NOISE COVARIANCE ESTIMATION FOR LINEAR SYSTEMS

Travis J. Arnold

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy
(Chemical Engineering)

at the
University of Wisconsin–Madison
2020

To my parents.

If I have seen further, it is by standing on the shoulders of giants.

Isaac Newton

Date of final oral examination: 7/20/2020

The dissertation is approved by the following members of the final oral committee:

James B. Rawlings, Professor and Mellichamp Process Control Chair, Chemical Engineering (University of California, Santa Barbara)

Christos T. Maravelias, Vilas Distinguished Achievement Professor and Paul A. Elfers Professor, Chemical and Biological Engineering

Victor M. Zavala, Baldovin-DaPra Associate Professor, Chemical and Biological Engineering

Michael D. Graham, Steenbock Professor of Engineering and Harvey D. Spangler Professor, Chemical and Biological Engineering

Robert D. Nowak, Keith and Jane Morgan Nosbusch Professor, Electrical and Computer Engineering

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Surely, I said, knowledge is the food of the soul.

PLATO

ABSTRACT

Dynamical systems are affected both by process noises, which affect the internal states of the system directly, and by measurement noises, the inherent error associated with measuring the system outputs. The subject of this dissertation, noise covariance estimation, is the task of identifying separately the process and measurement noise covariance matrices. Covariance estimation has received a great deal of attention in the literature in recent years, as algorithms designed for this task show great promise to play a key role within the hierarchy of control activities in industrial applications.

Previous work in this arena has often focused primarily on the development of covariance estimation algorithms without paying much mind to the important question of covariance matrix identifiability. We address this shortcoming, demonstrating that the conditions required for unique identification of the covariance matrices, along with the Kalman filter gain matrix, are relatively stringent compared to the unique identifiability conditions for the Kalman predictor gain matrix.

We study two different classes of covariance estimation algorithms: autocovariance least squares (ALS)-based methods and maximum likelihood (ML)-based methods. We present the first generally tractable means for calculation of the optimal weighting matrix for ALS problems, which enables the implementation of several novel ALS-based algorithms. We compare ALS-based and ML-based covariance estimation algorithms, both in terms of the statistical properties (namely, unbiasedness and consistency) which theory promises that they confer upon the covariance estimates, and in terms of empirically observed performance. In the examples studied, we observe that ML-based methods provide the highest quality covariance estimates. However, the computation time required for solving ML optimization problems scales dramatically with number of data points, so ML-based methods are usually intractable for all but small to moderately sized problems. In contrast, we find that ALS-based methods, where the weighting matrix is properly estimated, provide covariance estimates which have nearly the same quality rela-

tive to ML-based methods, but the computation time for ALS-based methods is independent of the number of data points.

Gratitude is not only the greatest of virtues, but the parent of all the others.

Cicero

ACKNOWLEDGMENTS

First and foremost, thanks are owed to my advisor, James B. Rawlings. I judge it a marvelous stroke of fortune that I found my way into his research group some five and a half years ago. Jim has challenged me every step of the way through graduate school, and I am a better researcher, thinker, and writer for it. Above all, I am grateful for the patience that Jim has often extended to me along the way; I am not sure that I would have finished this dissertation without this courtesy.

Others are due explicit recognition as well. Thanks to Christos Maravelias, Victor Zavala, Mike Graham, and Rob Nowak for serving on my dissertation committee, and thanks a second time to Professors Zavala and Rawlings for overseeing my TA experiences. More broadly, I am grateful to all the professors that I have interacted with during my time at UW, including those that taught my classes and those that served on my preliminary exam and fourth year meeting committees. I am also grateful to Quinn Zheng and Jacky Lao for arranging and overseeing my internship at AspenTech in Houston during summer 2016.

Raising a child takes a village, and the same is true of writing a dissertation. I am indebted to the University of Wisconsin–Madison as a whole, and the Chemical and Biological Engineering department in particular, for fostering an engaging and productive research environment. Thanks are especially owed to the CBE department staff. I will single out Kate Fanis, Russ Poyner, Theresa Neisius, and the since retired Mary Heimbecker and Kathy Heinzen as targets for my gratitude.

Working in the Rawlings group has been a privilege, and I am beholden to my colleagues who made coming to the office every day so enjoyable. I am the fourth student in our group to write a dissertation about noise covariance estimation, and I am grateful to Brian Odelson, Murali Rajamani, and Megan Zagrobelny for their work on the subject which gave me a firm foundation on which to begin my project. I am thankful for my friendships with Michael Risbeck, Nishith Patel, Doug Allan, Min Yao, Koty McAllister, Pratyush Kumar, Marco Vaccari, and Joel Andersson. Michael deserves a special mention, as his willingness to share his

seemingly encyclopedic knowledge of control theory, mathematics, programming, etc., was an invaluable resource. I will consider myself lucky if I enjoy the company of my coworkers in future roles half as much as I did in the Rawlings group. I wish the best of luck to Koty and Pratyush, as well as Chris Kuo-LeBlanc and Steven Kuntz, as they continue their studies in Santa Barbara.

All of the friends I made in the CBE department have contributed in a large manner to my experience in graduate school. I am grateful for all of the ChEGS events attended, faculty lunch leftovers consumed, ultimate frisbee games played in, etc., especially and particularly because of the people that I shared these experiences with. I cannot hope to provide a complete list of people who helped shape my time in Madison, but the following names deserve to be mentioned in addition to those which have already been given: Kevin Barnett, Joe Chada, Ryan Clark, Garrett Cogburn, Taylor Cook, Nat Eagan, Curran Gahan, Siddarth Krishna, Alec Linot, Parth Magnolia, Chris Mehrer, Frank Nguyen, and Jonathan Sheavly.

Rarely have I been bored during the last six years, and that is just as much due to my life away from campus as it is to my life on campus. I am not sure if one can thank a city or state, per se, but I want to express my gratitude for the city of Madison and the state of Wisconsin. I could not have asked for a better setting in which to live, work, and play during graduate school. I am grateful to Hoofers for helping me to get outside and enjoy all the natural beauty that Wisconsin has to offer, from the driftless region in the southwest to the northwoods abutting the shores of Lake Superior. I also am grateful to everyone I played with in the Old Buzzard Hockey League C division in Sun Prairie for ensuring that Friday night was the part of the week that I looked forward to the most. It is hard to put into words just how much the Madison community means to me; I will miss it terribly moving forward.

Dad, Mom, Alison, and all of my other family members. What can I say. I am so grateful for the constancy of their love and support for me. The stresses of graduate school have been all the more manageable knowing what a strong support system I have behind me. I am thankful to my dad in particular for encouraging me to pursue graduate education in the first place; I probably would not have considered a PhD if not for him. Thanks are also owed to Allie Kerr for the special brand of affection and confidence she has instilled me with over the past two years.

Travis Arnold
Madison, WI
July 2020

Mathematical discoveries, small or great, are never born of spontaneous generation.

Henri Poincaré

CHAPTER 0

NOTATION, USEFUL IDENTITIES, AND DEFINITIONS

This chapter provides a smörgåsbord of notation, facts, identities, and definitions that are used throughout the dissertation.

- $\mathbb{R}$ denotes the set of real numbers, $\mathbb{R}^n$ denotes the set of real $n \times 1$ vectors, and $\mathbb{R}^{m \times n}$ denotes the set of real $m \times n$ matrices.
- $\mathbb{C}$ denotes the set of complex numbers.
- LI and LD abbreviate linearly independent and linearly dependent, respectively.
- A^{T} is the transpose of matrix A.
- $A^{\mathrm{F}} = A + A^{\mathrm{T}}$ for square matrix A. This notation is not standard, but it is convenient for this dissertation, in Chapter 4 in particular.
- $0_{m,n}$ is the $m \times n$ zero matrix. The subscripts may be omitted if the dimensions are clear from context.
- $\emptyset$ is the empty set.
- $\implies$ means "implies", $\impliedby$ means "is implied by", and $\iff$ means "implies and is implied by", or "if and only if".

- I_n is the n-dimensional identity matrix. The subscript may be omitted if the dimension is clear from context.
- ι_n is the $n \times 1$ column vector of ones, e.g., $\iota_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.
- δ_{kj} is the Kronecker delta:

$$\delta_{kj} \coloneqq \begin{cases} 0, & k \neq j \\ 1, & k = j \end{cases}$$

- $\subseteq$ and $\supseteq$ denote subset and superset, respectively.
- $W^\perp$ denotes the orthogonal complement of subspace W. A useful fact is that $W \subseteq Y$ if and only if $Y^\perp \subseteq W^\perp$.
- $\operatorname{col} X$ denotes the column space of matrix X.
- $\operatorname{null} X$ denotes the null space of matrix X. A useful identity is $\operatorname{null} X^{\mathrm{T}} = (\operatorname{col} X)^\perp$.
- $\operatorname{eig} X$ denotes the set of all eigenvalues of square matrix X.
- $|X|$ denotes the determinant of matrix X (which is equivalent to absolute value if X is a scalar).
- $|||\cdot|||$ denotes a matrix norm and $\|\cdot\|$ denotes a vector norm. In particular, $|||\cdot|||_F$ is the Frobenius norm, $|||\cdot|||_\infty$ is the matrix infinity norm, $|||\cdot|||_p$ is the matrix p-norm, and $\|\cdot\|_p$ is the vector p-norm.
- $X > 0$ and $X \geq 0$ denote that square, symmetric matrix X is positive definite and positive semidefinite, respectively.
- $\|x\|_W^2 \coloneqq x^{\mathrm{T}} W x$ is the W-weighted inner product of vector x with itself. The weighting matrix W is generally taken to be positive semidefinite.
- X^+ denotes the Moore-Penrose inverse of matrix X, which is the unique matrix satisfying

$$X = XX^+X \qquad X^+ = X^+XX^+ \qquad XX^+ = (XX^+)^{\mathrm{T}} \qquad X^+X = (X^+X)^{\mathrm{T}}$$

- X_{s} denotes the vectorization operation, i.e., stacking the columns of matrix X into a vector. Some authors use $\operatorname{vec} X$ for this operation.

- X_{ss} denotes stacking only the elements on and below the diagonal of matrix X into a vector. This operator is most typically applied to symmetric matrices to extract the unique elements, but can also be applied to matrices with more rows than columns where the upper square block is symmetric for the same purpose. Some authors use vech X for this operation.

- $A \otimes B$ denotes the Kronecker product of matrices A and B.

- $(ABC)_{\mathrm{s}} = (C^{\mathrm{T}} \otimes A)B_{\mathrm{s}}$ for any conforming matrices A, B, and C.

- $\mathscr{D}_k$ is the duplication matrix, which is the unique $k^2 \times k(k+1)/2$ matrix such that $\mathscr{D}_k X_{\mathrm{ss}} = X_{\mathrm{s}}$ for all symmetric $k \times k$ matrices X. The matrix $\mathscr{D}_k$ has LI columns for all k, so $X_{\mathrm{ss}} = \mathscr{D}_k^{+} X_{\mathrm{s}}$ for all such X.

- $\mathscr{D}_{\ell,k}$ is the unique $\ell k \times \ell k - k(k-1)/2$ matrix such that $\mathscr{D}_{\ell,k} X_{\mathrm{ss}} = X_{\mathrm{s}}$ for all $\ell \times k$ matrices X with $\ell \geq k$ and the upper $k \times k$ block of X symmetric. The matrix $\mathscr{D}_{\ell,k}$ has LI columns for all $\ell \geq k$, so $X_{\mathrm{ss}} = \mathscr{D}_{\ell,k}^{+} X_{\mathrm{s}}$ for all such X. Moreover, $\mathscr{D}_{\ell,k}$ represents a generalization of the duplication matrix, and $\mathscr{D}_{k,k} = \mathscr{D}_k$.

- $\mathscr{K}_{\ell,k}$ is the commutation matrix, which is the unique $\ell k \times \ell k$ matrix such that $\mathscr{K}_{\ell,k} X_{\mathrm{s}} = (X^{\mathrm{T}})_{\mathrm{s}}$ for all $\ell \times k$ matrices X. The commutation matrix satisfies the identities $\mathscr{K}_{\ell,k}^{\mathrm{T}} = \mathscr{K}_{\ell,k}^{-1} = \mathscr{K}_{k,\ell}$. The notation $\mathscr{K}_k$ is used as shorthand for $\mathscr{K}_{k,k}$. The commutation matrix is so named because it allows the Kronecker product to "commute" in the following sense: for matrices A and B with dimensions $\ell \times k$ and $r \times q$ respectively,

$$\mathscr{K}_{r,\ell}(A \otimes B)\mathscr{K}_{k,q} = B \otimes A$$

- The singular value decomposition (SVD) of matrix X is

$$X = USV^{\mathrm{T}} = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^{\mathrm{T}} \\ V_2^{\mathrm{T}} \end{bmatrix} = U_1 \Sigma V_1^{\mathrm{T}}$$

  where U and V are orthogonal matrices and Σ is diagonal and positive definite. The entries along the diagonal of Σ are called the singular values of X. Moreover, the columns of U_1, U_2, V_1, and V_2 form bases for $\operatorname{col} X$, $\operatorname{null} X^{\mathrm{T}}$, $\operatorname{col} X^{\mathrm{T}}$, and $\operatorname{null} X$, respectively.

- (y_k) denotes a sequence, and $y_{0:m}$ denotes the finite sequence $(y_0, y_1, \cdots, y_m)$. The notation $y_k \to y$ indicates that the sequence (y_k) converges to y.

- $\mathrm{E}[x]$ denotes the expectation of random variable x.
- $\mathrm{cov}(x)$ denotes the covariance matrix of random variable x.
- In the context of probability, $x \mid y$ denotes that random variable x is conditioned on y. For example, $p_{x|y}(x \mid y)$ is the conditional probability density of x given y, and $\mathrm{E}[x \mid y]$ is the conditional expectation of x given y.
- In the context of probability, $x; \theta$ denotes that the distribution of random variable x is parameterized by θ. For example, for we can denote the probability density function of x by $p_x(x; \theta)$. Occasionally in this dissertation we omit displaying '$; \theta$', in places where it is expedient to simplify the notation.
- $x \sim \mathrm{N}(\mu, \Sigma)$ denotes that x is a (possibly multivariate) normal random variable with mean vector μ and covariance matrix Σ.
- $\mathrm{tr}\, A$ denotes the trace of square matrix A. A useful property is $\mathrm{tr}\, AB = \mathrm{tr}\, BA$ if both of the matrix products AB and BA exist and are square (note that the products do not need to be the same size). Another useful property is $\mathrm{tr}\, A^{\mathrm{T}} A = \sum_i \sigma_{A,i}^2$ where $\sigma_{A,i}$ denotes a singular value of A.
- $\mathrm{d}X$ denotes the differential of X. The differential operator is distinct from, although equivalent to, the derivative, but the differential is often less cumbersome to use with matrix calculus.

Definition 1. A square matrix A is *stable* if all of its eigenvalues are strictly inside the unit circle.

Remark 0.1. A useful property of stable matrices is that A stable implies $\lim_{k\to\infty} A^k = 0$ (Horn and Johnson, 2013, Theorem 3.2.5.2).

Definition 2. For $A \in \mathbb{R}^{n\times n}$ and $C \in \mathbb{R}^{p\times n}$, the pair (A, C) is *observable* if the *observability matrix*

$$\mathcal{O} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

has LI columns, i.e., if $\mathrm{rank}\, \mathcal{O} = n$. Similarly, the pair (A, B) is *controllable* if $(A^{\mathrm{T}}, B^{\mathrm{T}})$ is observable.

Definition 3. A linear time-invariant discrete time state space system is in *observability canonical form* if it is written as

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_{k+1} = \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_k \qquad y = \begin{bmatrix} C_1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_k$$

with (A_{11}, C_1) observable. Then x_1 denotes the observable states and x_2 the unobservable states. If A_{22} is a stable matrix, then the unobservable states are stable, and the system is said to be *detectable*. Similarly, the pair (A, B) is *stabilizable* if $(A^{\mathrm{T}}, B^{\mathrm{T}})$ is detectable.

Definition 4. If $\hat{X}$ is an estimate for X, the *bias* of $\hat{X}$ is the quantity $\mathrm{E}[\hat{X} - X]$. The estimate $\hat{X}$ is *unbiased* if $\mathrm{E}[\hat{X} - X] = 0$ for all possible values of X, and *biased* otherwise.

Definition 5. A sequence of scalar random variables (X_n) *converges in probability* to the random variable X if $\lim_{n\to\infty} \Pr(|X_n - X| \geq \varepsilon) = 0$ for all $\varepsilon > 0$. The notation $X_n \xrightarrow{P} X$ denotes convergence in probability.

Remark 0.2. For vector and matrix random variables, convergence in probability is defined similarly to Definition 5 except that the absolute value is replaced with any norm.

Definition 6. If $\hat{X}_{N_d}$ is an estimate for X, where the subscript N_d denotes the number of data points used to calculate $\hat{X}_{N_d}$, then $\hat{X}_{N_d}$ is *consistent* if $\hat{X}_{N_d} \xrightarrow{P} X$ as $N_d \to \infty$.

Remark 0.3. Definition 6 generally assumes the limit X does not depend on the number of data points, which is not always the case in this dissertation. When we say that $\hat{X}_{N_d}$ is consistent for X_{N_d}, we mean $\hat{X}_{N_d} - X_{N_d} \xrightarrow{P} 0$ as $N_d \to \infty$.

Definition 7. A random sequence (X_n) is *stationary* if all of its joint probability densities are time-invariant, i.e., if for arbitrary integers $m \geq 1$, t, and $k_j \geq 1$, $j \in \{1, 2, \cdots, m\}$,

$$p_{X_{k_1}, X_{k_2}, \cdots X_{k_m}}(x_{k_1}, x_{k_2}, \cdots x_{k_m}) = p_{X_{k_1+t}, X_{k_2+t}, \cdots X_{k_m+t}}(x_{k_1+t}, x_{k_2+t}, \cdots x_{k_m+t})$$

Definition 8. A random sequence (X_n) is *wide-sense stationary* if its first and second moments are time-invariant, i.e, if for arbitrary integers $k \geq 1$, $k_1 \geq 1$, $k_2 \geq 1$, and t,

$$\mathrm{E}[X_k] = \mathrm{E}[X_{k+t}]$$
$$\mathrm{E}[X_{k_1} X_{k_2}^{\mathrm{T}}] = \mathrm{E}[X_{k_1+t} X_{k_2+t}^{\mathrm{T}}]$$

Remark 0.4. Stationary implies wide-sense stationary. For Gaussian (i.e., normally distributed) random sequences, stationary and wide-sense stationary are equivalent.

See Magnus and Neudecker (2019), Horn and Johnson (2013), Graham and Rawlings (2013, Chapter 1), and Golub and Van Loan (2013) as general references on linear algebra. See Gubner (2006) and Graham and Rawlings (2013, Chapter 4) as general references on probability and statistics. See Rawlings et al. (2017) as a general reference on control related topics.

If you don't know where you are going, you might wind up someplace else.

Yogi Berra

CHAPTER 1

INTRODUCTION

This chapter is organized as follows:

- Section 1.1 provides background information on process control, process modeling, and model predictive control.
- Section 1.2 introduces the subject of noise covariance estimation, the primary motivating topic for the entire dissertation.
- Section 1.3 lists the three guiding research objectives of the dissertation.
- Section 1.4 gives an outline of the dissertation.

1.1 Background

1.1.1 Process control

Control engineering is the discipline which is tasked with developing and implementing strategies to operate dynamical systems under desirable operating conditions. The field is diverse; core ideas of control have been applied successfully in a great range of contexts. For example, control engineers play an integral role in industries such as chemical processing, aerospace, robotics, and computer hardware and software design, to name only a few. Few of the advances which have been

attained in the modern incarnations of these industries would be possible without the contributions of control engineers.

Process control is the subfield of control engineering concerned with continuous production processes. Such processes are ubiquitous in the chemical industries. As such, chemical engineers have played an integral role in the development of process control, both in theory and in application. Collaboration with practitioners in the chemical process industries has inspired to a large degree the work which comprises this dissertation. Despite the particular source of inspiration, the results herein are presented in a general context, and have potential to be useful in a wide scope of control applications beyond the process industries.

At a high level, the goal of process control is to operate a system in the way which produces the most possible benefits while respecting all relevant constraints. Benefits are generally economic, e.g., it is desirable to operate a chemical plant at a high feedthrough rate with minimal energy consumption in order to maximize profits. Constraints are sometimes economic, e.g., the changing prices of feedstocks can affect the optimal plant operating conditions, but can also be related to equipment and safety, e.g., plants are limited by the amount and nature of the physical equipment they have available, and all of that equipment must always be operated at safe temperature and pressure.

The activities which comprise process control operate on time scales which span a wide range. This concept is known as the hierarchy of process control; Figure 1.1 gives a visual demonstration. Let us briefly summarize the function of each level in the hierarchy:

- **Planning and scheduling.** Long term planning of plant operation.
- **System identification.** System identification is the task of identifying the system model from data. A wide variety of system identification techniques have been developed and are used widely in industrial applications. Ljung (1999) provides a good overview of the system identification literature. Subspace methods bear particular mention as a notable class of system identification techniques which have gained traction in recent years. Katayama (2005) and van Overschee and De Moor (1996) are two good references on subspace methods.

  The frequency with which system reidentification must occur depends on the time scale of process drift for the system. In general, complete system identification should be performed as infrequently as possible. System identification requires a data set for which all plant inputs are persistently excited

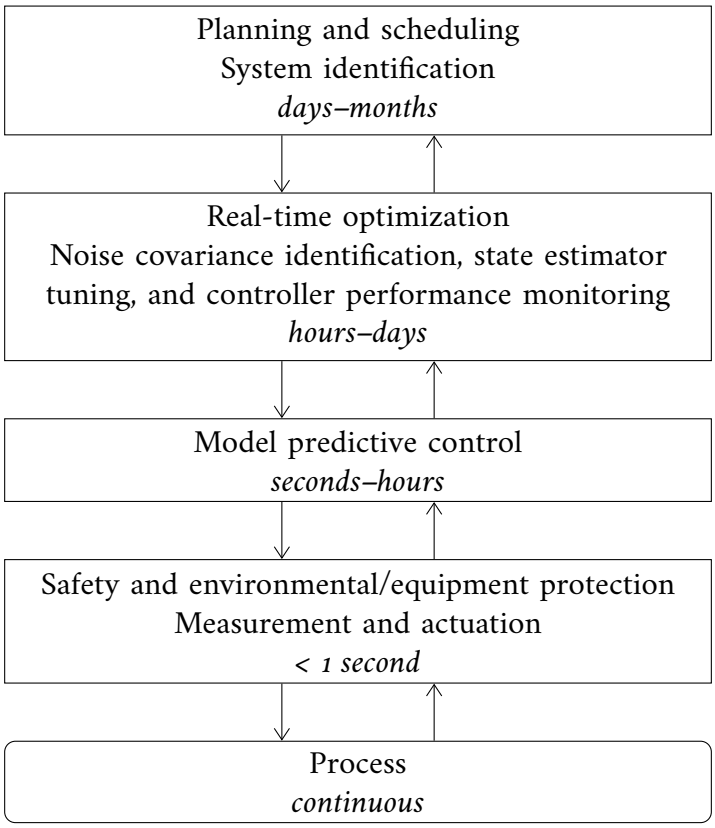

Figure 1.1: The hierarchy of process control. The typpical time scale for each level is included in italics. Varying versions of this figure are common in the process control literature; see, for example, Seborg et al. (2017, Figure 1.8, page 8).

so that the identification algorithm will properly account for the effect of all of the inputs on the system dynamics. Generating such data sets can be time consuming, costly, and disruptive to normal plant operation, since such data sets generally are not generated under normal plant operating conditions.

- **Real-time optimization.** Abbreviated RTO. This layer frequently adjusts the desired plant set point based on current plant conditions, process disturbances, and economic conditions.
- **Noise covariance identification, state estimator tuning, and controller performance monitoring.** These three activities are distinct, but they are closely related in that noise covariance identification is a task which en-

ables the applications of state estimator tuning and controller performance monitoring. Noise covariance identification is the primary subject of this dissertation as a whole, and we defer an overview of this task to Section 1.2.

- **Model predictive control.** Abbreviated MPC. This layer implements a control law which is responsible for driving the plant to the set point received from the RTO layer. We discuss in MPC more detail in Section 1.1.3.
- **Safety and environmental/equipment protection.** This layer refers to functions which are in place to always ensure the plant is operated safely. For example, alarm management systems are part of this layer.
- **Measurement and actuation.** This layer refers to hardware and instrumentation such as sensors, transmitters, and actuators such as control valves. Oftentimes this layer includes proportional-integral-derivative (PID) controllers which receive their set points from the MPC layer, a strategy which is known as cascade control (Ogunnaike and Ray, 1994, Section 16.1.1, pages 567–570).

Not all levels of the hierarchy are included in every process control application. In particular, planning and scheduling and RTO can be expensive to implement, and may be omitted depending on the particular needs of a plant. MPC is sometimes omitted as well, but in this situation a different control strategy, most often traditional PID control, must be implemented in its stead.

1.1.2 Process modeling

The modeling framework for this dissertation is the linear time-invariant discrete time state space dynamic model:

$$x_{k+1} = Ax_k + Bu_k + Gw_k \tag{1.1}$$

$$y_k = Cx_k + Du_k + Hv_k \tag{1.2}$$

At discrete time k, $x_k \in \mathbb{R}^n$ is the state, $y_k \in \mathbb{R}^p$ is the output, $u_k \in \mathbb{R}^m$ is the input, $w_k \in \mathbb{R}^g$ is the process noise, and $v_k \in \mathbb{R}^h$ is the measurement noise. The matrices A, B, C, D, G, and H have conforming dimensions: $A \in \mathbb{R}^{n\times n}$ is the state transition matrix, $B \in \mathbb{R}^{n\times m}$ is the input shaping matrix, $C \in \mathbb{R}^{p\times n}$ is the state measurement matrix, $D \in \mathbb{R}^{p\times m}$ is the input feedthrough matrix, and $G \in \mathbb{R}^{n\times g}$ and $H \in \mathbb{R}^{p\times h}$

are the process and measurement noise shaping matrices, respectively. We also encounter systems without inputs, in which case the model is

$$x_{k+1} = Ax_k + Gw_k \tag{1.3}$$
$$y_k = Cx_k + Hv_k \tag{1.4}$$

We refer to (1.1) and (1.2) together as the process system with inputs and (1.3) and (1.4) together as the process system without inputs.

In this dissertation, we do not ever place any restrictions on the noise shaping matrices G and H, but there is generally not any reason to consider the case where either G or H has LD columns. For example, if G has LD columns, one can use $\tilde{G}$ in its stead where the columns of $\tilde{G}$ form a basis for the column space of G, and hence are LI.

The noises are modeled as random variables. The first and second moments are

$$\mathrm{E}\begin{bmatrix} w_k \\ v_\ell \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \mathrm{E}\left[\begin{bmatrix} w_k \\ v_\ell \end{bmatrix}\begin{bmatrix} w_i^{\mathrm{T}} & v_j^{\mathrm{T}} \end{bmatrix}\right] = \begin{bmatrix} Q\delta_{ki} & S\delta_{kj} \\ S^{\mathrm{T}}\delta_{\ell i} & R\delta_{\ell j} \end{bmatrix}$$

In Chapters 2 and 3, no particular probability distribution is assumed for the noises, but in Chapters 4, 5 and 7 and parts of Chapter 6, we assume that the noises are normally distributed.

It is common that the number of statistically independent process noises is less than the number of states, i.e., $g < n$. If one has an idea of which states are affected by independent process noises, one may design the process noise shaping matrix G to reflect this knowledge. Similarly, one can conceive of situations where the number of statistically independent measurement noises is less than the number of outputs, i.e. $h < p$. This case is less common than the $g < n$ case, but an example of when it might occur is if several output signals are sent through the same multiplexing device. For this reason we also consider the general measurement noise shaping matrix H instead of assuming $H = I_p$.

There exist types of dynamic models other than state space models, most notably, transfer function models. Dynamic models can generally be converted from one type to another, but we have chosen to use the state space framework because it is the most natural fit for the topic of interest in this dissertation, and because state space models have become the most common choice throughout the modern control literature.

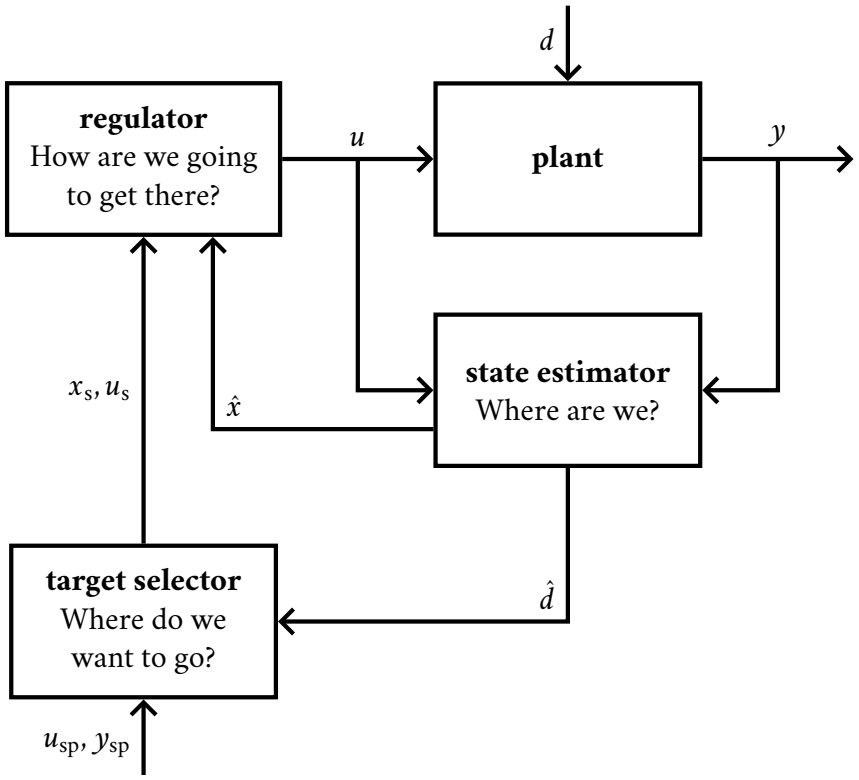

Figure 1.2: A typical model predictive control algorithm.

1.1.3 Model predictive control

In the last three decades, model predictive control (MPC) has become a leading paradigm in the field of advanced process control. MPC is a framework for building a control algorithm rather than a particular control algorithm itself. MPC has been extended in many directions as authors have used it for different kinds of systems and applications. For example, there is linear MPC, nonlinear MPC, distributed MPC, and explicit MPC, to name just a few different flavors. In this regard, the MPC literature is not monolithic. The unifying aspect of MPC algorithms is that at every sampling time, an optimal control problem, which incorporates the system model, is solved and the optimal inputs are applied to the plant.

Although MPC is diverse, we still find it useful to discuss the structure of a typical MPC feedback loop. As shown in Figure 1.2, there are three main components, which have the following functions:

- **State estimator.** Based on the inputs which enter the plant and the measurements which are taken from it, the state estimator returns an estimate for the current state of the plant, $\hat{x}$. If disturbances d are affecting the plant, the state estimator may be designed to also provide a disturbance estimate, $\hat{d}$. For linear systems, the optimal state estimator is the Kalman filter, which is presented in Section 2.1. The Kalman filter is also known as the linear quadratic estimator (LQE).

- **Target selector.** Let $(u_{\text{sp}}, y_{\text{sp}})$ denote the desired input/output operating condition for the plant, perhaps determined by RTO. Then, using $(u_{\text{sp}}, y_{\text{sp}})$ and the state and disturbance estimates, $\hat{x}$ and $\hat{d}$, the function of the target selector is to return a steady state/input pair $(x_{\text{s}}, u_{\text{s}})$ that the regulator should drive the plant towards. One generally cannot pass $(u_{\text{sp}}, y_{\text{sp}})$ directly into the regulator, because the optimization problem that the regulator solves is usually in terms of x and u, not u and y.

- **Regulator.** The regulator determines the inputs u which should be applied to the plant to drive it towards the steady state $(x_{\text{s}}, u_{\text{s}})$ generated by the target selector. For linear systems, the most canonical choice for the regulator is the linear quadratic regulator (LQR).

When authors refer to a "control problem", they most often mean a regulation problem, but the state estimation and target selection components are just as important to the MPC algorithm as a whole.

Industrial MPC applications have lagged significantly behind theoretical advances. Despite this lag, MPC remains popular both as a study of academic inquiry and among industrial practitioners. Two of the biggest reasons for this popularity are that MPC usually provides superior control performance for interacting multivariate systems compared to traditional control strategies such as multiloop PID, and that MPC provides a natural way to impose constraints upon the control problem. Rawlings et al. (2017) is a monograph describing the state of the art of MPC theory. Qin and Badgwell (2003) give a survey of industrial MPC implementations as of 2003. Lahiri (2017) provides a more contemporary history, overview, and survey of industrial MPC technology.

1.2 Motivation: noise covariance estimation

This dissertation is concerned with the covariance estimation problem, i.e., estimating the matrices Q, R, and S from plant data, assuming the model matrices A, B, C, D, G, and H are known. Knowledge of Q, R, and S is useful for constructing a state estimator with optimal performance, such as the Kalman predictor and filter[1] or a moving horizon estimator, and for controller performance monitoring, as described by Zagrobelny et al. (2013).

Covariance estimation can be considered to be a subtopic of system identification. One might question the usefulness of covariance estimation relative to system identification, given that the scope of the latter subsumes that of the former. However, the narrower scope of covariance estimation is precisely why it is useful and relevant for many applications. As is mentioned in Section 1.1.1, full system identification is often disruptive, costly, and time consuming because it requires the generation of a particular data set where all of the system inputs are persistently excited. In contrast, the noises are naturally exciting, so any set of input-output data can be used for covariance estimation. Furthermore, the statistics of the random disturbances affecting a process generally change in nature more frequently than do the deterministic components of the dynamics. Thus, in practice it is usually useful to perform covariance reidentification on a faster time scale than full system reidentification, which is why covariance identification is a level below system identification in Figure 1.1, the hierarchy of process control.

The majority of the covariance estimation literature has focused on developing estimation algorithms with relatively little attention paid to the topic of identifiability. In this dissertation, we devote considerable effort in both of these directions. Regarding the identifiability topic, the system identification literature frequently, and often tacitly, references the fact that for linear systems, in general it is possible to identify the Kalman predictor form of the system but not the process form. Rarely, however, does one find any discussion about the identifiability of the Kalman filter. One example is Tse and Anton (1972), who discuss the identifiability of parameters in a general context and apply their results to identifying linear systems with normally distributed noises. Similarly, subspace identification methods generally provide an estimate for the Kalman predictor gain, but not for the Kalman filter gain (Qin, 2006).

[1] Many authors do not distinguish in terminology between prediction problems and filtering problems, using the term filter for both. However the distinction matters for our purposes, and we will use the two terms accordingly.

In this dissertation, we consider two general classes of covariance estimation algorithms:

- **Correlation methods** infer Q, R, and S from the sample autocovariance of the outputs. Mehra (1970) was the first to use this strategy. Another correlation method is autocovariance least squares (ALS) (Odelson et al., 2006b; Rajamani and Rawlings, 2009). ALS is the technique considered in this dissertation. ALS improves upon Mehra's method by reducing the number of steps in the estimation procedure from three to one, which has the benefit of reducing the variance of the final covariance estimates.
- **Maximum likelihood (ML) methods** estimate the covariance matrices by maximizing the likelihood function of the output data. The ML problem is often solved using the expectation maximization (EM) algorithm (Dempster et al., 1977).

There are classes of covariance estimation methods beyond correlation and ML methods, such as covariance matching and Bayesian methods. However, such methods are beyond the scope of this dissertation. The covariance estimation literature is large, and it is not our goal to provide an exhaustive history of said literature. We cite many relevant works throughout the dissertation as appropriate, and we refer the reader to Duník et al. (2017) and Odelson (2003, Chapter 3) for comprehensive literature reviews of the subject.

1.3 Research objectives

The research presented in this dissertation serves three primary objectives:

Objective 1. Obtain precise results regarding the identifiability of the covariance matrices (Q, R, S), the Kalman predictor gain matrix K, and the Kalman filter gain matrix L.

Objective 2. Improve the ALS technique for covariance estimation by developing a tractable way to calculate and estimate the optimal weighting matrix for the ALS problem.

Objective 3. Compare the ALS and ML approaches to covariance estimation, both in terms of the statistical properties which are theoretically guaranteed by various algorithms and in terms of empirically observed performance of the algorithms.

These objectives are somewhat ambiguous as stated. Additional context and the details of how we work toward these objectives may of course be found within the main body of the dissertation. We revisit these objectives in the conclusion, examining to what extent they are met, and exploring what avenues are fertile for future research.

1.4 Outline and summary of dissertation

The following is an outline of the dissertation, briefly summarizing the contributions of each chapter:

- Chapter 0 gives mathematical notation, as well as useful identities and definitions which are used throughout the dissertation.
- Chapter 1 (this chapter) introduces the dissertation, giving background information, motivation, and stating the three main research objectives.
- Chapter 2 presents the Kalman predictor and filter, derives the $\mathcal{K}$-error system, and demonstrates two different parameterizations for the steady-state autocovariance of both the system outputs and the $\mathcal{K}$-innovations.
- Chapter 3 derives the ALS technique for covariance estimation, gives precise uniqueness conditions for the ALS problem under the assumptions $G = I_n$ and $H = I_p$, and discusses some of the implications that the results from Chapter 2 have for ALS problems.
- Chapter 4 develops the first generally tractable method to calculate the optimal weighting matrix for the ALS problem and compares the method with previous work from other researchers.
- Chapter 5 discusses the ML technique for covariance estimation, including the expectation maximization (EM) algorithm for solving ML problems.
- Chapter 6 studies the theoretical properties, namely unbiasedness and consistency, which are theoretically guaranteed by various ALS-based and ML-based covariance estimation algorithms.
- Chapter 7 presents numerical examples as a basis for emprical comparison of ALS-based and ML-based covariance estimation algorithms.

- Chapter 8 concludes the dissertation, summarizing the extent to which each of the three main research objectives are accomplished and discussing potential avenues for future research.

- Appendix A discusses the linear regression model, the resulting minimum variance affine unbiased estimator (MVAUE), and the relationship of the MVAUE to least squares problems.

Chapters 2 and 3 serve Objective 1, Chapter 4 serves Objective 2, and Chapters 6 and 7 serve Objective 3. Chapter 5 does not directly serve any of the three objectives, but the information it provides is a necessary precursor to the material in Chapters 6 and 7.

The material in Chapters 2 and 3 has been adapted and extended from two previous publications, the first of which is Arnold and Rawlings (2018). That paper considers only the process system without inputs, (1.3) and (1.4), and it assumes that the matrix A is stable, $G = I_n$, and $H = I_p$. The next paper, Arnold and Rawlings (2020a), allows the system to have inputs, weakens the A stable assumption to (A, C) detectable, and drops the $G = I_n$, and $H = I_p$ assumptions altogether. The version presented here, in Chapters 2 and 3, is mostly similar to Arnold and Rawlings (2020a), but features additional results, discussion, and more detailed versions of proofs. Similarly, the material in Chapters 4–7 is an extended version of Arnold and Rawlings (2020b).

A place for everything and everything in its place.

Benjamin Franklin

CHAPTER 2

THE KALMAN PREDICTOR AND FILTER AND AUTOCOVARIANCE PARAMETERIZATIONS

This chapter serves Objective 1 of the dissertation, which is to obtain precise results regarding the identifiability of the covariance matrices (Q, R, S), the Kalman predictor gain matrix K, and the Kalman filter gain matrix L. Motivated by results from Anderson and Moore (1979, Section 9.2), we show that in general, K is uniquely identifiable from the output autocovariance under less stringent conditions than are either (Q, R, S) or L.

Our techniques require that the outputs sequence be wide-sense stationary, which is not the case if the system has inputs or if A is not stable. The extension comes by processing the outputs with a stable (but possibly non-optimal) Kalman predictor with constant gain matrix $\mathcal{K}$. This procedure generates what we call the $\mathcal{K}$-innovations sequence, which is the output to what we call the $\mathcal{K}$-error system. The $\mathcal{K}$-innovations sequence is wide-sense stationary and thus amenable to analysis.

Throughout this chapter and the next, results are given for two cases: first for the process system without inputs and A stable and then the analogous result for the $\mathcal{K}$-error system. As such, these two chapters are fairly repetitive. In fact, the abridged version of the results given in Arnold and Rawlings (2020a) includes only the $\mathcal{K}$-error system track for the sake of brevity. Here, however, both tracks are provided. Although the results are wholly analogous for the two systems, there are

subtle differences in the derivations and proofs between the systems, so we include both tracks for the purpose of endowing the literature with a complete exposition.

This chapter is organized as follows:

- Section 2.1 gives the Kalman predictor and filter equations and derives the error system and the $\mathcal{K}$-error system.
- Section 2.2 gives an overview of steady-state process statistics for linear systems. In particular it is shown that the output autocovariance can parameterized either by (Q, R, S) or by a different set of parameters, which we call Φ and Θ (or $\mathbb{\Phi}$ and $\mathbb{\Theta}$ for the $\mathcal{K}$-error system).
- Section 2.3 shows how to calculate the Kalman predictor gain K from (Φ, Θ) and from $(\mathbb{\Phi}, \mathbb{\Theta})$.
- Section 2.4 discusses the relationship between (Q, R, S) and (Φ, Θ) and between (Q, R, S) and $(\mathbb{\Phi}, \mathbb{\Theta})$. It is explained how assuming $S = 0$ changes the results.

The chapter also contains two appendices: 2.A and 2.B.

2.1 The Kalman predictor and filter

The Kalman predictor and filter estimates the state of the system based on the outputs. There are two different types of state estimate:

$$\hat{x}_{k|k-1} \coloneqq \mathrm{E}[x_k \mid y_{0:k-1}] \qquad \hat{x}_{k|k} \coloneqq \mathrm{E}[x_k \mid y_{0:k}]$$

The quantity $\hat{x}_{k|k-1}$ is called a predicted estimate and $\hat{x}_{k|k}$ is called a filtered estimate. The covariances of the estimates are

$$\begin{aligned} \Sigma_{k|k-1} &\coloneqq \mathrm{E}[(x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^{\mathrm{T}}] \\ \Sigma_{k|k} &\coloneqq \mathrm{E}[(x_k - \hat{x}_{k|k})(x_k - \hat{x}_{k|k})^{\mathrm{T}}] \end{aligned}$$

If x_0 is a random variable that is independent of (w_k, v_k) for all k, then $\hat{x}_{k|k-1}$, $\hat{x}_{k|k}$, $\Sigma_{k|k-1}$, and $\Sigma_{k|k}$ may be calculated recursively:

- Predictor form:

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k-1} + Bu_k + K_k(y_k - C\hat{x}_{k|k-1} - Du_k) \tag{2.1}$$
$$\Sigma_{k+1|k} = A\Sigma_{k|k-1}A^{\mathrm{T}} - K_k(A\Sigma_{k|k-1}C^{\mathrm{T}} + GSH^{\mathrm{T}})^{\mathrm{T}} + GQG^{\mathrm{T}} \tag{2.2}$$
$$K_k \coloneqq (A\Sigma_{k|k-1}C^{\mathrm{T}} + GSH^{\mathrm{T}})(C\Sigma_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+}$$

  Equation (2.2) is known as a discrete Riccati iteration.

- Filter form:

  - Time update:

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k} + Bu_k + GSH^{\mathrm{T}}(HRH^{\mathrm{T}})^{+}(y_k - C\hat{x}_{k|k} - Du_k)$$
$$\Sigma_{k+1|k} = [A - GSH^{\mathrm{T}}(HRH^{\mathrm{T}})^{+}C]\Sigma_{k|k}[A - GSH^{\mathrm{T}}(HRH^{\mathrm{T}})^{+}C]^{\mathrm{T}} + GQG^{\mathrm{T}} - GSH^{\mathrm{T}}(HRH^{\mathrm{T}})^{+}HS^{\mathrm{T}}G^{\mathrm{T}}$$

  - Measurement update:

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + L_k(y_k - C\hat{x}_{k|k-1} - Du_k)$$
$$\Sigma_{k|k} = \Sigma_{k|k-1} - L_kC\Sigma_{k|k-1}$$
$$L_k \coloneqq \Sigma_{k|k-1}C^{\mathrm{T}}(C\Sigma_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+}$$

See Anderson and Moore (1979, Sections 5.4 and 5.5) for derivations of the Kalman predictor and filter equations.[1] We call K_k the predictor gain and L_k the filter gain.

The estimates $\hat{x}_{k|k-1}$ and $\hat{x}_{k|k}$ are the minimum variance affine unbiased estimates of x_k from $y_{0:k-1}$ and $y_{0:k}$, respectively. If (w_k), (v_k), and x_0 are jointly Gaussian, then the result is stronger. In this case, $\hat{x}_{k|k-1}$ and $\hat{x}_{k|k}$ are the minimum variance unbiased estimates of x_k from $y_{0:k-1}$ and $y_{0:k}$, respectively. See Anderson and Moore (1979, Chapters 2 and 5) for discussion of these points.

Let μ_k and P_k denote the state mean and covariance, respectively, at time k:

$$\mu_k \coloneqq \mathrm{E}[x_k] \qquad P_k \coloneqq \mathrm{E}[(x_k - \mu_k)(x_k - \mu_k)^{\mathrm{T}}]$$

These quantities evolve as

$$\mu_{k+1} = A\mu_k + Bu_k \qquad P_{k+1} = AP_kA^{\mathrm{T}} + GQG^{\mathrm{T}}$$

[1] Anderson and Moore assume $H = I_p$, R invertible, and $D = 0$. The Kalman predictor and filter equations that shown here are valid if any combination of these three assumptions do not hold.

Furthermore, μ_0 and P_0 are used to initialize the Kalman predictor:

$$\hat{x}_{0|-1} = \mu_0 \qquad \Sigma_{0|-1} = P_0$$

We make the following assumptions throughout this dissertation:

Assumption 1. The Kalman predictor and filter is asymptotically time-invariant, i.e., $\Sigma_{k|k-1} \to \Sigma$ for some limit Σ.

Assumption 2. The limit Σ is independent of the initial condition $\Sigma_{0|-1}$ (provided that $\Sigma_{0|-1} \geq 0$).

Assumption 3. For any initial condition $\Sigma_{0|-1} \geq 0$,

$$\operatorname{rank}(C\Sigma_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}}) \to \operatorname{rank}(C\Sigma C^{\mathrm{T}} + HRH^{\mathrm{T}})$$

Assumptions 1 and 2 cannot be guaranteed without a priori knowledge of Q, R, and S, but they hold in most practical cases. Anderson and Moore (1979, Sections 4.4 and 5.4) and Silverman (1976) discuss sufficient conditions for Assumptions 1 and 2 to hold. Assumption 3 is a necessary and sufficient condition for

$$(C\Sigma_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+} \to (C\Sigma C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+}$$

(Stewart, 1969), which implies

$$K_k \to K \coloneqq (A\Sigma C^{\mathrm{T}} + GSH^{\mathrm{T}})(C\Sigma C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+}$$
$$L_k \to L \coloneqq \Sigma C^{\mathrm{T}}(C\Sigma C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+}$$

and that Σ satisfies a steady-state version of (2.2), known as discrete algebraic Riccati equation (DARE):

$$\Sigma = A\Sigma A^{\mathrm{T}} - K(A\Sigma C^{\mathrm{T}} + GSH^{\mathrm{T}})^{\mathrm{T}} + GQG^{\mathrm{T}} \tag{2.3}$$

Assumption 3 holds if (but not necessarily only if) $HRH^{\mathrm{T}} > 0$, as this condition implies $C\Sigma_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}} > 0$ for all k. This case is the most typical one encountered in practice, and many authors assume $HRH^{\mathrm{T}} > 0$ when presenting the Kalman predictor and filter. In this scenario, inverse may replace Moore-Penrose inverse in the Kalman predictor and filter equations.

We define the functions that map (Q, R, S) to K and L as

$$f_K(Q, R, S) \coloneqq (A\Sigma C^{\mathrm{T}} + GSH^{\mathrm{T}})(C\Sigma C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+}$$
$$f_L(Q, R, S) \coloneqq \Sigma C^{\mathrm{T}}(C\Sigma C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+}$$

where Σ is the unique solution to (2.3). The domain of both f_K and f_L is taken to be the set of all (Q, R, S) such that

$$\begin{bmatrix} Q & S \\ S^{\mathrm{T}} & R \end{bmatrix} \geq 0$$

We also define

$$\tilde{f}_K(Q, R) \coloneqq f_K(Q, R, 0) \qquad \tilde{f}_L(Q, R) \coloneqq f_L(Q, R, 0)$$

The domain of both $\tilde{f}_K$ and $\tilde{f}_L$ is taken to be the set of all (Q, R) such that $Q \geq 0$ and $R \geq 0$.

2.1.1 The error system

If A is unstable and/or the system has inputs, (x_k) and (y_k) are not guaranteed to be wide-sense stationary as $k \to \infty$. Another approach must be used to generate a sequence that is asymptotically wide-sense stationary. To this end, we define the predictor estimate error e_k and the innovation z_k:

$$e_k \coloneqq x_k - \hat{x}_{k|k-1} \qquad z_k \coloneqq y_k - C\hat{x}_{k|k-1} - Du_k$$

These quantities evolve as a time-varying linear system:

$$\begin{aligned} e_{k+1} &= x_{k+1} - \hat{x}_{k+1|k} \\ &= Ax_k + Bu_k + Gw_k - A\hat{x}_{k|k-1} - Bu_k - K_k(y_k - C\hat{x}_{k|k-1} - Du_k) \\ &= Ax_k + Gw_k - A\hat{x}_{k|k-1} - K_k(Cx_k - C\hat{x}_{k|k-1} + Hv_k) \\ &= (A - K_kC)(x_k - \hat{x}_{k|k-1}) + Gw_k - K_kHv_k \\ &= \bar{A}_k e_k + \bar{G}_k \bar{w}_k \end{aligned} \tag{2.4}$$

$$\begin{aligned} z_k &= y_k - C\hat{x}_{k|k-1} - Du_k \\ &= Cx_k + Du_k + Hv_k - C\hat{x}_{k|k-1} - Du_k \\ &= C(x_k - \hat{x}_{k|k-1}) + Hv_k \\ &= Ce_k + Hv_k \end{aligned} \tag{2.5}$$

$$\bar{A}_k \coloneqq A - K_kC \qquad \bar{G}_k \coloneqq \begin{bmatrix} G & -K_kH \end{bmatrix} \qquad \bar{w}_k \coloneqq \begin{bmatrix} w_k \\ v_k \end{bmatrix}$$

We refer to (2.4) and (2.5) together as the error system. For convenience, we define

$$\bar{Q} \coloneqq \mathrm{E}[\bar{w}_k \bar{w}_k^{\mathrm{T}}] = \begin{bmatrix} Q & S \\ S^{\mathrm{T}} & R \end{bmatrix} \qquad \bar{S} \coloneqq \mathrm{E}[\bar{w}_k v_k^{\mathrm{T}}] = \begin{bmatrix} S \\ R \end{bmatrix}$$

The Kalman predictor and filter applied to the error system yields

$$
\begin{aligned}
\hat{e}_{k+1|k} &:= \mathrm{E}[e_{k+1} \mid z_{0:k}] \\
&= \bar{A}_k \hat{e}_{k|k-1} + \bar{K}_k (z_k - C\hat{e}_{k|k-1}) \\
&= \bar{A}_k \hat{e}_{k|k} + \bar{G}_k \bar{S} R^+ H^+ (z_k - C\hat{e}_{k|k}) \\
\hat{e}_{k+1|k+1} &:= \mathrm{E}[e_{k+1} \mid z_{0:k+1}] = \hat{e}_{k+1|k} + \bar{L}_{k+1}(z_{k+1} - C\hat{e}_{k+1|k}) \\
\bar{\Sigma}_{k+1|k} &:= \mathrm{E}[(e_{k+1} - \hat{e}_{k+1|k})(e_{k+1} - \hat{e}_{k+1|k})^{\mathrm{T}}] \\
&= \bar{A}_k \bar{\Sigma}_{k|k-1} \bar{A}_k^{\mathrm{T}} - \bar{K}_k (\bar{A}_k \bar{\Sigma}_{k|k-1} C^{\mathrm{T}} + \bar{G}_k \bar{S} H^{\mathrm{T}})^{\mathrm{T}} + \bar{G}_k \bar{Q} \bar{G}_k^{\mathrm{T}} \\
&= (\bar{A}_k - \bar{G}_k \bar{S} R^+ H^+ C) \bar{\Sigma}_{k|k} (\bar{A}_k - \bar{G}_k \bar{S} R^+ H^+ C)^{\mathrm{T}} + \bar{G}_k (\bar{Q} - \bar{S} R^+ \bar{S}^{\mathrm{T}}) \bar{G}_k^{\mathrm{T}} \\
\bar{\Sigma}_{k|k} &:= \mathrm{E}[(e_k - \hat{e}_{k|k})(e_k - \hat{e}_{k|k})^{\mathrm{T}}] = \bar{\Sigma}_{k|k-1} - \bar{L}_k C \bar{\Sigma}_{k|k-1} \\
\bar{K}_k &:= (\bar{A}_k \bar{\Sigma}_{k|k-1} C^{\mathrm{T}} + \bar{G}_k \bar{S} H^{\mathrm{T}})(C \bar{\Sigma}_{k|k-1} C^{\mathrm{T}} + HRH^{\mathrm{T}})^+ \\
\bar{L}_k &:= \bar{\Sigma}_{k|k-1} C^{\mathrm{T}} (C \bar{\Sigma}_{k|k-1} C^{\mathrm{T}} + HRH^{\mathrm{T}})^+ \\
\bar{\mu}_{k+1} &:= \mathrm{E}[e_{k+1}] = \bar{A}_k \bar{\mu}_k \\
\bar{P}_{k+1} &:= \mathrm{E}[(e_{k+1} - \bar{\mu}_{k+1})(e_{k+1} - \bar{\mu}_{k+1})^{\mathrm{T}}] = \bar{A}_k \bar{P}_k \bar{A}_k^{\mathrm{T}} + \bar{G}_k \bar{Q} \bar{G}_k^{\mathrm{T}} \\
s_k &:= z_k - C\hat{e}_{k|k-1}
\end{aligned}
$$

Theorem 2.1. *The Kalman properties for the error system are related to those for the process system as follows:*

$$
\begin{aligned}
\hat{e}_{k|k-1} &= 0 & \bar{\Sigma}_{k|k} &= \Sigma_{k|k} & \bar{\mu}_k &= 0 \\
\hat{e}_{k|k} &= \hat{x}_{k|k} - \hat{x}_{k|k-1} & \bar{K}_k &= 0 \quad (2.7) & \bar{P}_k &= \Sigma_{k|k-1} \quad (2.9) \\
\bar{\Sigma}_{k|k-1} &= \Sigma_{k|k-1} \quad (2.6) & \bar{L}_k &= L_k \quad (2.8) & s_k &= z_k
\end{aligned}
$$

Proof. First, using linearity of conditional expectation (Gubner, 2006, Example 13.22) gives $\hat{e}_{k|k-1} = \hat{x}_{k|k-1} - \hat{x}_{k|k-1} = 0$ and $\hat{e}_{k|k} = \hat{x}_{k|k} - \hat{x}_{k|k-1}$. Next, we get

$$
\begin{aligned}
\bar{\Sigma}_{k|k-1} &= \mathrm{E}[(e_k - \hat{e}_{k|k-1})(e_k - \hat{e}_{k|k-1})^{\mathrm{T}}] \\
&= \mathrm{E}[e_k e_k^{\mathrm{T}}] = \mathrm{E}[(x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^{\mathrm{T}}] = \Sigma_{k|k-1}
\end{aligned}
$$

Similarly, $e_k - \hat{e}_{k|k} = x_k - \hat{x}_{k|k}$, so

$$
\bar{\Sigma}_{k|k} = \mathrm{E}[(e_k - \hat{e}_{k|k})(e_k - \hat{e}_{k|k})^{\mathrm{T}}] = \mathrm{E}[(x_k - \hat{x}_{k|k})(x_k - \hat{x}_{k|k})^{\mathrm{T}}] = \Sigma_{k|k}
$$

Then

$$
\begin{aligned}
\bar{K}_k &= (\bar{A}_k \bar{\Sigma}_{k|k-1} C^{\mathrm{T}} + \bar{G}_k \bar{S} H^{\mathrm{T}})(C \bar{\Sigma}_{k|k-1} C^{\mathrm{T}} + HRH^{\mathrm{T}})^+ \\
&= [A\Sigma_{k|k-1} C^{\mathrm{T}} + GSH^{\mathrm{T}} - K_k (C\Sigma_{k|k-1} C^{\mathrm{T}} + HRH^{\mathrm{T}})](C\Sigma_{k|k-1} C^{\mathrm{T}} + HRH^{\mathrm{T}})^+
\end{aligned}
$$

$$
\begin{aligned}
&= (A\Sigma_{k|k-1}C^{\mathrm{T}} + GSH^{\mathrm{T}})(C\Sigma_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+} \\
&\quad - K_k(C\Sigma_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}})(C\Sigma_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+} \\
&= K_k - (A\Sigma_{k|k-1}C^{\mathrm{T}} + GSH^{\mathrm{T}})(C\Sigma_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+} \\
&\qquad\qquad \cdot (C\Sigma_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}})(C\Sigma_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+} \\
&= K_k - (A\Sigma_{k|k-1}C^{\mathrm{T}} + GSH^{\mathrm{T}})(C\Sigma_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+} \\
&= K_k - K_k \\
&= 0
\end{aligned}
$$

Moreover,

$$
\begin{aligned}
\bar{L}_k &= \bar{\Sigma}_{k|k-1}C^{\mathrm{T}}(C\bar{\Sigma}_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+} \\
&= \Sigma_{k|k-1}C^{\mathrm{T}}(C\Sigma_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+} \\
&= L_k
\end{aligned}
$$

Furthermore, because $\hat{x}_{k|k-1}$ is an unbiased estimator for x_k, we have $\bar{\mu}_k = \mathrm{E}[e_k] = \mathrm{E}[x_k - \hat{x}_{k|k-1}] = 0$. Thus $\bar{P}_k$ is equal to $\Sigma_{k|k-1}$:

$$
\bar{P}_k = \mathrm{E}[(e_k - \bar{\mu}_k)(e_k - \bar{\mu}_k)^{\mathrm{T}}] = \mathrm{E}[e_k e_k^{\mathrm{T}}] = \Sigma_{k|k-1}
$$

Finally, $s_k = z_k - C\hat{e}_{k|k-1} = z_k$. ■

Corollary 2.1. *The sequences* $(\bar{\Sigma}_{k|k-1})$, $(\bar{K}_k)$, $(\bar{L}_k)$, *and* $(\bar{P}_k)$ *converge to the following limits:*

$$
\bar{\Sigma}_{k|k-1} \to \bar{\Sigma} \coloneqq \Sigma \qquad \bar{K}_k \to \bar{K} \coloneqq 0 \qquad \bar{L}_k \to \bar{L} \coloneqq L \qquad \bar{P}_k \to \bar{P} \coloneqq \Sigma
$$

Proof. These results follow by taking limits of (2.6)–(2.9). ■

Because $K_k \to K$, we have $\bar{A}_k \to \bar{A} \coloneqq A - KC$ and $\bar{G}_k \to \bar{G} \coloneqq \begin{bmatrix} G & -KH \end{bmatrix}$. Thus the error system is time-invariant as $k \to \infty$:

$$
e_{k+1} = \bar{A}e_k + \bar{G}\bar{w}_k \qquad z_k = Ce_k + Hv_k
$$

Assumption 4. $\bar{A}$ is stable.

Assumption 4 holds in most practical cases; Anderson and Moore (1979, Section 5.4) and Silverman (1976) give sufficient conditions for it to hold. If it does hold, then (e_k) and (z_k) are wide-sense stationary as $k \to \infty$.

The Kalman predictor and filter of the error system is not particularly interesting on its own. After the outputs have been processed with the Kalman predictor and filter, additional information cannot be obtained by applying the Kalman predictor and filter on the innovations. These results have been given so they may be extended: our next task is to present a suboptimal version of the error system for which the Kalman predictor and filter are important.

2.1.2 The $\mathcal{K}$-error system

To repeat, if Assumption 4 holds, then innovations sequence (z_k) is asymptotically wide-sense stationary. The issue is that calculation of (z_k) is predicated on knowledge of (Q, R, S), which is the unknown information sought in the first place when solving a covariance estimation problem. Therefore a different strategy must be employed to generate a wide-sense stationary sequence for analysis.[2] To this end, we require another assumption:

Assumption 5. The system (A, C) is detectable, and given $\mathcal{K} \in \mathbb{R}^{n\times p}$, the matrix $\mathbb{A} := A - \mathcal{K}C$ is stable.[3]

Using (2.1) with $K_k = \mathcal{K}$ for all k generates the $\mathcal{K}$-predicted estimates $\hat{x}_{k|k-1}$, the $\mathcal{K}$-predictor estimate errors $\mathbb{e}_k$, and the $\mathcal{K}$-innovations $\mathbb{z}_k$:

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k-1} + Bu_k + \mathcal{K}(y_k - C\hat{x}_{k|k-1} - Du_k)$$
$$\mathbb{e}_k := x_k - \hat{x}_{k|k-1} \qquad \mathbb{z}_k := y_k - C\hat{x}_{k|k-1} - Du_k$$

The quantities $\mathbb{e}_k$ and $\mathbb{z}_k$ evolve as a time-invariant linear system

$$\mathbb{e}_{k+1} = \mathbb{A}\mathbb{e}_k + \mathbb{G}\bar{w}_k \tag{2.10}$$
$$\mathbb{z}_k = C\mathbb{e}_k + Hv_k \tag{2.11}$$

where $\mathbb{G} := \begin{bmatrix} G & -\mathcal{K}H \end{bmatrix}$. Equations (2.10) and (2.11) are derived in the same manner as are (2.4) and (2.5). We refer to (2.10) and (2.11) together as the $\mathcal{K}$-error system. The

[2] In addition to the method described in this section of passing to the $\mathcal{K}$-innovations, Section 2.B gives an alternate method to transform the outputs into a wide-sense stationary sequence that may be employed if the system has inputs but A is stable.

[3] Detectability is necessary and sufficient for such a $\mathcal{K}$ to exist (Anderson and Moore, 1979, Appendix C).

Kalman predictor and filter equations applied to the $\mathcal{K}$-error system are

$$\begin{aligned}
\hat{\mathbb{e}}_{k+1|k} &\coloneqq \mathrm{E}[\mathbb{e}_{k+1} \mid \mathbb{z}_{0:k}] \\
&= \mathbb{A}\hat{\mathbb{e}}_{k|k-1} + \mathbb{K}_k(\mathbb{z}_k - C\hat{\mathbb{e}}_{k|k-1}) \\
&= \mathbb{A}\hat{\mathbb{e}}_{k|k} + \mathbb{G}\bar{S}R^{+}H^{+}(\mathbb{z}_k - C\hat{\mathbb{e}}_{k|k}) \\
\hat{\mathbb{e}}_{k+1|k+1} &\coloneqq \mathrm{E}[\mathbb{e}_{k+1} \mid \mathbb{z}_{0:k+1}] = \hat{\mathbb{e}}_{k+1|k} + \mathbb{L}_{k+1}(\mathbb{z}_{k+1} - C\hat{\mathbb{e}}_{k+1|k}) \\
\mathbb{\Sigma}_{k+1|k} &\coloneqq \mathrm{E}[(\mathbb{e}_{k+1} - \hat{\mathbb{e}}_{k+1|k})(\mathbb{e}_{k+1} - \hat{\mathbb{e}}_{k+1|k})^{\mathrm{T}}] \\
&= \mathbb{A}\mathbb{\Sigma}_{k|k-1}\mathbb{A}^{\mathrm{T}} - \mathbb{K}_k(\mathbb{A}\mathbb{\Sigma}_{k|k-1}C^{\mathrm{T}} + \mathbb{G}\bar{S}H^{\mathrm{T}})^{\mathrm{T}} + \mathbb{G}\bar{Q}\mathbb{G}^{\mathrm{T}} \\
&= (\mathbb{A} - \mathbb{G}\bar{S}R^{+}H^{+}C)\mathbb{\Sigma}_{k|k}(\mathbb{A} - \mathbb{G}\bar{S}R^{+}H^{+}C)^{\mathrm{T}} + \mathbb{G}(\bar{Q} - \bar{S}R^{+}\bar{S}^{\mathrm{T}})\mathbb{G}^{\mathrm{T}} \\
\mathbb{\Sigma}_{k|k} &\coloneqq \mathrm{E}[(\mathbb{e}_k - \hat{\mathbb{e}}_{k|k})(\mathbb{e}_k - \hat{\mathbb{e}}_{k|k})^{\mathrm{T}}] = \mathbb{\Sigma}_{k|k-1} - \mathbb{L}_k C\mathbb{\Sigma}_{k|k-1} \\
\mathbb{K}_k &\coloneqq (\mathbb{A}\mathbb{\Sigma}_{k|k-1}C^{\mathrm{T}} + \mathbb{G}\bar{S}H^{\mathrm{T}})(C\mathbb{\Sigma}_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+} \\
\mathbb{L}_k &\coloneqq \mathbb{\Sigma}_{k|k-1}C^{\mathrm{T}}(C\mathbb{\Sigma}_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+} \\
\mathbb{\mu}_{k+1} &\coloneqq \mathrm{E}[\mathbb{e}_{k+1}] = \mathbb{A}\mathbb{\mu}_k \\
\mathbb{P}_{k+1} &\coloneqq \mathrm{E}[(\mathbb{e}_{k+1} - \mathbb{\mu}_{k+1})(\mathbb{e}_{k+1} - \mathbb{\mu}_{k+1})^{\mathrm{T}}] = \mathbb{A}\mathbb{P}_k\mathbb{A}^{\mathrm{T}} + \mathbb{G}\bar{Q}\mathbb{G}^{\mathrm{T}} \qquad (2.12) \\
\mathbb{s}_k &\coloneqq \mathbb{z}_k - C\hat{\mathbb{e}}_{k|k-1}
\end{aligned}$$

We require an assumption on the initial condition for the $\mathcal{K}$-error system:

Assumption 6. The estimate $\hat{\mathbb{x}}_{0|-1}$ is unbiased for x_0; that is, $\hat{\mathbb{x}}_{0|-1} = \mu_0$.

The $\mathcal{K}$-error system, (2.10) and (2.11), and the error system, (2.4) and (2.5), are structurally identical, with the only difference being that that $\mathcal{K}$ replaces K_k in the former. Next we give results analogous to those in Theorem 2.1:

Theorem 2.2. *The Kalman properties for the $\mathcal{K}$-error system are related to those for the process system as follows:*

$$\hat{\mathbb{e}}_{k|k-1} = \hat{x}_{k|k-1} - \hat{\mathbb{x}}_{k|k-1} \qquad \mathbb{\Sigma}_{k|k-1} = \Sigma_{k|k-1} \qquad (2.13) \qquad \mathbb{\mu}_k = 0 \qquad (2.14)$$

$$\hat{\mathbb{e}}_{k|k} = \hat{x}_{k|k} - \hat{\mathbb{x}}_{k|k-1} \qquad \mathbb{\Sigma}_{k|k} = \Sigma_{k|k} \qquad \mathbb{s}_k = z_k$$

$$\mathbb{K}_k = K_k - \mathcal{K}(C\Sigma_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}})(C\Sigma_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+} \qquad (2.15)$$

$$\mathbb{L}_k = L_k \qquad (2.16)$$

$$\mathbb{P}_k = \mathrm{E}[(x_k - \mathbb{x}_{k|k-1})(x_k - \mathbb{x}_{k|k-1})^{\mathrm{T}}] \qquad (2.17)$$

Proof. First, using linearity of conditional expectation (Gubner, 2006, Example 13.22) gives $\hat{\mathbb{e}}_{k|k-1} = \hat{x}_{k|k-1} - \hat{\mathbb{x}}_{k|k-1}$ and $\hat{\mathbb{e}}_{k|k} = \hat{x}_{k|k} - \hat{\mathbb{x}}_{k|k-1}$. Next, noting that

$\mathbb{e}_k - \hat{\mathbb{e}}_{k|k-1} = x_k - \hat{x}_{k|k-1}$, we get

$$\begin{aligned}\mathbb{\Sigma}_{k|k-1} &= \mathrm{E}[(\mathbb{e}_k - \hat{\mathbb{e}}_{k|k-1})(\mathbb{e}_k - \hat{\mathbb{e}}_{k|k-1})^{\mathrm{T}}] \\ &= \mathrm{E}[(x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^{\mathrm{T}}] = \Sigma_{k|k-1}\end{aligned}$$

Similarly, $\mathbb{e}_k - \hat{\mathbb{e}}_{k|k} = x_k - \hat{x}_{k|k}$, so

$$\mathbb{\Sigma}_{k|k} = \mathrm{E}[(\mathbb{e}_k - \hat{\mathbb{e}}_{k|k})(\mathbb{e}_k - \hat{\mathbb{e}}_{k|k})^{\mathrm{T}}] = \mathrm{E}[(x_k - \hat{x}_{k|k})(x_k - \hat{x}_{k|k})^{\mathrm{T}}] = \Sigma_{k|k}$$

Then

$$\begin{aligned}\mathbb{K}_k &= (\mathbb{A}\mathbb{\Sigma}_{k|k-1}C^{\mathrm{T}} + \mathbb{G}\bar{S}H^{\mathrm{T}})(C\mathbb{\Sigma}_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+} \\ &= [(A - \mathcal{K}C)\Sigma_{k|k-1}C^{\mathrm{T}} + GSH^{\mathrm{T}} - \mathcal{K}HRH^{\mathrm{T}}](C\Sigma_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+} \\ &= (A\Sigma_{k|k-1}C^{\mathrm{T}} + GSH^{\mathrm{T}})(C\Sigma_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+} \\ &\quad - \mathcal{K}(C\Sigma_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}})(C\Sigma_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+} \\ &= K_k - \mathcal{K}(C\Sigma_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}})(C\Sigma_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+}\end{aligned}$$

Moreover,

$$\begin{aligned}\mathbb{L}_k &= \mathbb{\Sigma}_{k|k-1}C^{\mathrm{T}}(C\mathbb{\Sigma}_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+} \\ &= \Sigma_{k|k-1}C^{\mathrm{T}}(C\Sigma_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+} \\ &= L_k\end{aligned}$$

Furthermore, from Assumption 6 we have $\mathbb{\mu}_0 = \mathrm{E}[\mathbb{e}_0] = \mathrm{E}[x_0 - \hat{x}_{0|-1}] = 0$. Then, since $\mathbb{\mu}_{k+1} = \mathbb{A}\mathbb{\mu}_k$, by induction $\mathbb{\mu}_k = 0$ for all $k \geq 1$. Then

$$\mathbb{P}_k = \mathrm{E}[(\mathbb{e}_k - \mathbb{\mu}_k)(\mathbb{e}_k - \mathbb{\mu}_k)^{\mathrm{T}}] = \mathrm{E}[\mathbb{e}_k\mathbb{e}_k^{\mathrm{T}}] = \mathrm{E}[(x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^{\mathrm{T}}]$$

Finally,

$$\begin{aligned}\mathbb{s}_k &= \mathbb{z}_k - C\hat{\mathbb{e}}_{k|k-1} \\ &= y_k - C\hat{x}_{k|k-1} - Du_k - C(\hat{x}_{k|k-1} - \hat{x}_{k|k-1}) \\ &= y_k - C\hat{x}_{k|k-1} - Du_k \\ &= z_k\end{aligned}$$

■

Remark 2.1. The matrix $\mathbb{P}_0$ satisfies $\mathbb{P}_0 = \mathrm{E}[(x_0 - \mu_0)(x_0 - \mu_0)^{\mathrm{T}}] = P_0 = \Sigma_{0|-1}$.

Remark 2.2. Sometimes μ_0 is not known a priori, in which case Assumption 6 may not hold. All of the results of Theorem 2.2 still apply in this situation, except for (2.14) and (2.17). However, because $\mathbb{A}$ is stable, (2.14) and (2.17) hold as $k \to \infty$.

Corollary 2.2. *The sequences* $(\mathbb{\Sigma}_{k|k-1})$, $(\mathbb{K}_k)$, $(\mathbb{L}_k)$, *and* $(\mathbb{P}_k)$ *converge to the following limits:*

$$\mathbb{\Sigma}_{k|k-1} \to \mathbb{\Sigma} := \Sigma \qquad \mathbb{L}_k \to \mathbb{L} := L \qquad \mathbb{P}_k \to \mathbb{P}$$
$$\mathbb{K}_k \to \mathbb{K} := K - \mathcal{K}(C\Sigma C^{\mathrm{T}} + HRH^{\mathrm{T}})(C\Sigma C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+}$$

where $\mathbb{P}$ *is the unique solution to the discrete Lyapunov equation* $\mathbb{P} = \mathbb{A}\mathbb{P}\mathbb{A}^{\mathrm{T}} + \mathbb{G}\bar{Q}\mathbb{G}^{\mathrm{T}}$.

Proof. These results follow by taking limits of (2.12), (2.13), (2.15), and (2.16). The results for $\mathbb{P}_k$ and $\mathbb{K}_k$ also rely on Corollary 2.9 (in Appendix 2.A) and Assumption 3, respectively. ■

2.2 Steady-state process statistics

Next, we display the steady state autocovariance for the process system, the error system, and the $\mathcal{K}$-error system.

2.2.1 Process system

Consider the process system without inputs, (1.3) and (1.4). Assume A is stable. Then, applying Corollary 2.9 from Appendix 2.A, as $k \to \infty$, $\mu_k \to 0$ and $P_k \to P$ where P is the unique solution to the discrete Lyapunov equation

$$P = APA^{\mathrm{T}} + GQG^{\mathrm{T}} \tag{2.18}$$

The map $P \leftrightarrow GQG^{\mathrm{T}}$ is bijective, which can be seen by vectorizing (2.18) and rearranging:

$$P_{\mathrm{s}} = F(GQG^{\mathrm{T}})_{\mathrm{s}} = F(G \otimes G)Q_{\mathrm{s}} \tag{2.19}$$
$$F := [I_{n^2} - (A \otimes A)]^{-1}$$

Stability of A guarantees that F is invertible and Q symmetric guarantees that P is symmetric (see Corollaries 2.8 and 2.9 in Appendix 2.A). As $k \to \infty$, the random processes (x_k) and (y_k) are wide-sense stationary with zero mean and the following autocovariances:

$$\mathrm{E}[x_{k+j}x_k^{\mathrm{T}}] = A^j P, \quad j \geq 0 \tag{2.20}$$
$$\mathrm{E}[y_k y_k^{\mathrm{T}}] = \Phi \tag{2.21}$$
$$\mathrm{E}[y_{k+j}y_k^{\mathrm{T}}] = CA^{j-1}\Theta, \quad j \geq 1 \tag{2.22}$$

$$\Phi := CPC^{\mathrm{T}} + HRH^{\mathrm{T}} \tag{2.23}$$

$$\Theta := APC^{\mathrm{T}} + GSH^{\mathrm{T}} \tag{2.24}$$

See Anderson and Moore (1979, Section 4.3) for derivations of these results.

2.2.2 Error system

We now present analogous results for the error system, (2.4) and (2.5). We already showed $\bar{P}_k \to \bar{P} = \Sigma$ (see Theorem 2.1). Additionally, again using Corollary 2.9, $\bar{P}$ is the unique solution to the discrete Lyapunov equation

$$\begin{aligned}\bar{P} &= \bar{A}\bar{P}\bar{A}^{\mathrm{T}} + \bar{G}\bar{Q}\bar{G}^{\mathrm{T}} \\ &= \bar{A}\bar{P}\bar{A}^{\mathrm{T}} + GQG^{\mathrm{T}} + KHRH^{\mathrm{T}}K^{\mathrm{T}} - GSH^{\mathrm{T}}K^{\mathrm{T}} - KHS^{\mathrm{T}}G^{\mathrm{T}}\end{aligned} \tag{2.25}$$

Because $\bar{P} = \Sigma$, (2.25) can be considered an alternative way to write the DARE (2.3). Moreover, the map $\bar{P} \leftrightarrow GQG^{\mathrm{T}} + KHRH^{\mathrm{T}}K^{\mathrm{T}} - GSH^{\mathrm{T}}K^{\mathrm{T}} - KHS^{\mathrm{T}}G^{\mathrm{T}}$ is bijective, which can be seen by vectorizing (2.25) and rearranging:

$$\begin{aligned}\bar{P}_{\mathrm{s}} &= \bar{F}(GQG^{\mathrm{T}} + KHRH^{\mathrm{T}}K^{\mathrm{T}} - GSH^{\mathrm{T}}K^{\mathrm{T}} - KHS^{\mathrm{T}}G^{\mathrm{T}})_{\mathrm{s}} \\ &= \bar{F}[(G \otimes G)Q_{\mathrm{s}} + (KH \otimes KH)R_{\mathrm{s}} - (KH \otimes G)S_{\mathrm{s}} - (G \otimes KH)(S^{\mathrm{T}})_{\mathrm{s}}] \\ &= \bar{F}[(G \otimes G)Q_{\mathrm{s}} + (KH \otimes KH)R_{\mathrm{s}} - (KH \otimes G)S_{\mathrm{s}} - (G \otimes KH)\mathscr{K}_{g,h}S_{\mathrm{s}}] \\ &= \bar{F}[(G \otimes G)Q_{\mathrm{s}} + (KH \otimes KH)R_{\mathrm{s}} - (KH \otimes G)S_{\mathrm{s}} - \mathscr{K}_n(KH \otimes G)S_{\mathrm{s}}] \\ &= \bar{F}[(G \otimes G)Q_{\mathrm{s}} + (KH \otimes KH)R_{\mathrm{s}} - (I_{n^2} + \mathscr{K}_n)(KH \otimes G)S_{\mathrm{s}}] \\ \bar{F} &:= [I_{n^2} - (\bar{A} \otimes \bar{A})]^{-1}\end{aligned}$$

Stability of $\bar{A}$ (Assumption 4) guarantees that $\bar{F}$ is invertible and $\bar{Q}$ symmetric guarantees that $\bar{P}$ is symmetric (see Corollaries 2.8 and 2.9). As $k \to \infty$, the random processes (e_k) and (z_k) are wide-sense stationary with zero mean and the following autocovariances:

$$\begin{aligned}\mathrm{E}[e_{k+j}e_k^{\mathrm{T}}] &= \bar{A}^j\bar{P}, \quad j \geq 0 \\ \mathrm{E}[z_k z_k^{\mathrm{T}}] &= \bar{\Phi} \\ \mathrm{E}[z_{k+j}z_k^{\mathrm{T}}] &= C\bar{A}^{j-1}\bar{\Theta}, \quad j \geq 1 \\ \bar{\Phi} &:= C\bar{P}C^{\mathrm{T}} + HRH^{\mathrm{T}} \\ \bar{\Theta} &:= \bar{A}\bar{P}C^{\mathrm{T}} + \bar{G}\bar{S}H^{\mathrm{T}} \\ &= A\bar{P}C^{\mathrm{T}} + GSH^{\mathrm{T}} - K(C\bar{P}C^{\mathrm{T}} + HRH^{\mathrm{T}})\end{aligned}$$

2.2.3 $\mathcal{K}$-error system

Next we give analogous results for the $\mathcal{K}$-error system.

$$\begin{aligned}\mathbb{P} &= \mathbb{A}\mathbb{P}\mathbb{A}^{\mathrm{T}} + \mathbb{G}\bar{Q}\mathbb{G}^{\mathrm{T}} \\ &= \mathbb{A}\mathbb{P}\mathbb{A}^{\mathrm{T}} + GQG^{\mathrm{T}} + \mathcal{K}HRH^{\mathrm{T}}\mathcal{K}^{\mathrm{T}} - GSH^{\mathrm{T}}\mathcal{K}^{\mathrm{T}} - \mathcal{K}HS^{\mathrm{T}}G^{\mathrm{T}} \qquad (2.26)\\ \mathbb{P}_{\mathrm{s}} &= \mathbb{F}[(G\otimes G)Q_{\mathrm{s}} + (\mathcal{K}H\otimes\mathcal{K}H)R_{\mathrm{s}} - (I_{n^2} + \mathcal{K}_n)(\mathcal{K}H\otimes G)S_{\mathrm{s}}] \qquad (2.27)\\ \mathbb{F} &:= [I_{n^2} - (\mathbb{A}\otimes\mathbb{A})]^{-1}\end{aligned}$$

Stability of $\mathbb{A}$ guarantees that $\mathbb{F}$ is invertible and $\bar{Q}$ symmetric guarantees that $\mathbb{P}$ is symmetric (see Corollaries 2.8 and 2.9). As $k \to \infty$,

$$\begin{aligned}\mathrm{E}[\mathbb{e}_{k+j}\mathbb{e}_k^{\mathrm{T}}] &= \mathbb{A}^j\mathbb{P}, \quad j \geq 0\\ \mathrm{E}[\mathbb{z}_k\mathbb{z}_k^{\mathrm{T}}] &= \Phi \qquad (2.28)\\ \mathrm{E}[\mathbb{z}_{k+j}\mathbb{z}_k^{\mathrm{T}}] &= C\mathbb{A}^{j-1}\Theta, \quad j \geq 1 \qquad (2.29)\\ \Phi &:= C\mathbb{P}C^{\mathrm{T}} + HRH^{\mathrm{T}} \qquad (2.30)\\ \Theta &:= \mathbb{A}\mathbb{P}C^{\mathrm{T}} + \mathbb{G}\bar{S}H^{\mathrm{T}}\\ &= A\mathbb{P}C^{\mathrm{T}} + GSH^{\mathrm{T}} - \mathcal{K}(C\mathbb{P}C^{\mathrm{T}} + HRH^{\mathrm{T}}) \qquad (2.31)\end{aligned}$$

2.3 Calculating the Kalman predictor gain

In this section, we show how to calculate K from (Φ, Θ) (and, equivalently, $\bar{K}$ from $(\bar{\Phi}, \bar{\Theta})$ and $\mathbb{K}$ from (Φ, Θ)). We also show how to calculate $(\mathbb{K}, \Phi, \Theta)$ from (Φ, Θ) and (K, Φ, Θ) from (Φ, Θ). We remark that $\bar{K}$ and $\mathbb{K}$ are generally not as useful to know as is K. Nevertheless, results for $\bar{K}$ and $\mathbb{K}$ are included in this section for the sake of completeness.

2.3.1 Calculating K from (Φ, Θ)

For the process system, it is clear that knowing Q, R, and S is sufficient to calculate Σ, K, and L by recursively calculating the sequences $(\Sigma_{k+1|k})$, (K_k), and (L_k) until convergence. Less obvious is the fact that knowing Φ and Θ is sufficient to calculate K and $P - \Sigma$:

Theorem 2.3 (Anderson and Moore (1979), Theorem 9.2.1, pages 228–229)**.** *Consider the sequence* (T_k)*, initialized by* $T_0 = 0_{n\times n}$ *and defined recursively by*

$$T_{k+1} = AT_kA^{\mathrm{T}} + (AT_kC^{\mathrm{T}} - \Theta_k)(\Phi_k - CT_kC^{\mathrm{T}})^{+}(AT_kC^{\mathrm{T}} - \Theta_k)^{\mathrm{T}} \tag{2.32}$$

where

$$\Phi_k := CP_kC^{\mathrm{T}} + HRH^{\mathrm{T}} \tag{2.33}$$

$$\Theta_k := AP_kC^{\mathrm{T}} + GSH^{\mathrm{T}} \tag{2.34}$$

Then

$$T_k = P_k - \Sigma_{k|k-1} \tag{2.35}$$

$$K_k = -(AT_kC^{\mathrm{T}} - \Theta_k)(\Phi_k - CT_kC^{\mathrm{T}})^{+} \tag{2.36}$$

Proof. We use induction to prove (2.35) holds for all k. We have already established that it holds for $k = 0$: $\Sigma_{0|-1} = P_0$. We assume $P_k = \Sigma_{k|k-1} + T_k$ and show that this condition implies $P_{k+1} = \Sigma_{k+1|k} + T_{k+1}$:

$$\begin{aligned}
T_{k+1} &= AT_kA^{\mathrm{T}} + (AT_kC^{\mathrm{T}} - \Theta_k)(\Phi_k - CT_kC^{\mathrm{T}})^{+}(AT_kC^{\mathrm{T}} - \Theta_k)^{\mathrm{T}} \\
&= AP_kA^{\mathrm{T}} - A\Sigma_{k|k-1}A^{\mathrm{T}} \\
&\quad + (AP_kC^{\mathrm{T}} - A\Sigma_{k|k-1}C^{\mathrm{T}} - AP_kC^{\mathrm{T}} - GSH^{\mathrm{T}}) \\
&\qquad \cdot (CP_kC^{\mathrm{T}} + HRH^{\mathrm{T}} - CT_kC^{\mathrm{T}})^{+} \\
&\qquad \cdot (AP_kC^{\mathrm{T}} - A\Sigma_{k|k-1}C^{\mathrm{T}} - AP_kC^{\mathrm{T}} - GSH^{\mathrm{T}})^{\mathrm{T}} \\
&= AP_kA^{\mathrm{T}} - A\Sigma_{k|k-1}A^{\mathrm{T}} \\
&\quad + (A\Sigma_{k|k-1}C^{\mathrm{T}} - GSH^{\mathrm{T}})(C\Sigma_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+}(A\Sigma_{k|k-1}C^{\mathrm{T}} - GSH^{\mathrm{T}})^{\mathrm{T}} \\
&= AP_kA^{\mathrm{T}} + GQG^{\mathrm{T}} \\
&\quad - [A\Sigma_{k|k-1}A^{\mathrm{T}} + GQG^{\mathrm{T}} - (A\Sigma_{k|k-1}C^{\mathrm{T}} - GSH^{\mathrm{T}}) \\
&\qquad \cdot (C\Sigma_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+}(A\Sigma_{k|k-1}C^{\mathrm{T}} - GSH^{\mathrm{T}})^{\mathrm{T}}] \\
&= P_{k+1} - \Sigma_{k+1|k}
\end{aligned}$$

It is straightforward to verify that (2.36) holds for all k:

$$\begin{aligned}
K_k &= (A\Sigma_{k|k-1}C^{\mathrm{T}} + GSH^{\mathrm{T}})(C\Sigma_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+} \\
&= (AP_kC^{\mathrm{T}} - AT_kC^{\mathrm{T}} + GSH^{\mathrm{T}})(CP_kC^{\mathrm{T}} - CT_kC^{\mathrm{T}} + HRH^{\mathrm{T}})^{+} \\
&= (\Theta_k - AT_kC^{\mathrm{T}})(\Phi_k - CT_kC^{\mathrm{T}})^{+} \\
&= -(AT_kC^{\mathrm{T}} - \Theta_k)(\Phi_k - CT_kC^{\mathrm{T}})^{+}
\end{aligned}$$

■

Theorem 2.4. *The matrix T_k satisfies*

$$T_k = \mathrm{E}[(\hat{x}_{k|k-1} - \mu_k)(\hat{x}_{k|k-1} - \mu_k)^{\mathrm{T}}]$$

Proof. First, observe that

$$\mathrm{E}[(x_k - \hat{x}_{k|k-1})\mu_k^{\mathrm{T}}] = \mathrm{E}[(x_k - \hat{x}_{k|k-1})]\mu_k^{\mathrm{T}} = 0$$

Next, note that $\hat{x}_{k|k-1}$ is a linear function of the outputs. Thus, a consequence of the orthogonality principle/projection theorem (Anderson and Moore, 1979, Section 5.2) is that

$$\mathrm{E}[(x_k - \hat{x}_{k|k-1})\hat{x}_{k|k-1}^{\mathrm{T}}] = 0$$

We use these equations to obtain

$$\begin{aligned}
P_k &= \mathrm{E}[(x_k - \mu_k)(x_k - \mu_k)^{\mathrm{T}}] \\
&= \mathrm{E}[(x_k - \hat{x}_{k|k-1} + \hat{x}_{k|k-1} - \mu_k)(x_k - \hat{x}_{k|k-1} + \hat{x}_{k|k-1} - \mu_k)^{\mathrm{T}}] \\
&= \mathrm{E}[(x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^{\mathrm{T}}] + \mathrm{E}[(\hat{x}_{k|k-1} - \mu_k)(\hat{x}_{k|k-1} - \mu_k)^{\mathrm{T}}] \\
&= \Sigma_{k|k-1} + \mathrm{E}[(\hat{x}_{k|k-1} - \mu_k)(\hat{x}_{k|k-1} - \mu_k)^{\mathrm{T}}]
\end{aligned}$$

Therefore

$$T_k = P_k - \Sigma_{k|k-1} = \mathrm{E}[(\hat{x}_{k|k-1} - \mu_k)(\hat{x}_{k|k-1} - \mu_k)^{\mathrm{T}}] \qquad \blacksquare$$

Since $P_k \longrightarrow P$ and $\Sigma_{k|k-1} \longrightarrow \Sigma$, (2.33)–(2.35) imply that $\Phi_k \longrightarrow \Phi$, $\Theta_k \longrightarrow \Theta$, and $T_k \longrightarrow T := P - \Sigma$, respectively. Note that

$$\begin{aligned}
\Phi_k - CT_kC^{\mathrm{T}} &= CP_kC^{\mathrm{T}} + HRH^{\mathrm{T}} - CT_kC^{\mathrm{T}} \\
&= C(P_k - T_k)C^{\mathrm{T}} + HRH^{\mathrm{T}} \\
&= C\Sigma_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}}
\end{aligned}$$

Therefore, in light of Assumption 3, taking the limits of (2.32) and (2.36) implies

$$\begin{aligned}
T &= ATA^{\mathrm{T}} - K(ATC^{\mathrm{T}} - \Theta)^{\mathrm{T}} \\
K &= -(ATC^{\mathrm{T}} - \Theta)(\Phi - CTC^{\mathrm{T}})^{+}
\end{aligned}$$

Observe that it is necessary to know Q, R, S, and P_0 in order to calculate the full sequences (Φ_k) and (Θ_k). However, the steady-state Kalman predictor gain K can be still be calculated if only the limits Φ and Θ are known: use (2.32) and (2.36) with $T_0 = 0_{n\times n}$ and $\Phi_k = \Phi$ and $\Theta_k = \Theta$ for all k to calculate (T_k) and (K_k) until (K_k) converges.

2.3.2 Calculating $\bar{K}$ from $(\bar{\Phi}, \bar{\Theta})$

The results of Section 2.3.1 extend to the error system:

$$
\begin{aligned}
\bar{T}_0 &= 0_{n\times n} \\
\bar{T}_{k+1} &= \bar{A}_k \bar{T}_k \bar{A}_k^{\mathrm{T}} + (\bar{A}_k \bar{T}_k C^{\mathrm{T}} - \bar{\Theta}_k)(\bar{\Phi}_k - C\bar{T}_k C^{\mathrm{T}})^{+}(\bar{A}_k \bar{T}_k C^{\mathrm{T}} - \bar{\Theta}_k)^{\mathrm{T}} \\
&= \mathrm{E}[(\hat{e}_{k+1|k} - \bar{\mu}_{k+1})(\hat{e}_{k+1|k} - \bar{\mu}_{k+1})^{\mathrm{T}}] \\
\bar{\Phi}_k &:= C\bar{P}_k C^{\mathrm{T}} + HRH^{\mathrm{T}} \to \bar{\Phi} \\
\bar{\Theta}_k &:= \bar{A}_k \bar{P}_k C^{\mathrm{T}} + \bar{G}_k \bar{S} H^{\mathrm{T}} \to \bar{\Theta} \\
\bar{T}_k &= \bar{P}_k - \bar{\Sigma}_{k|k-1} \to \bar{T} := \bar{P} - \bar{\Sigma} = \bar{A}\bar{T}\bar{A}^{\mathrm{T}} - \bar{K}(\bar{A}\bar{T}C^{\mathrm{T}} - \bar{\Theta})^{\mathrm{T}} \\
\bar{K}_k &= -(\bar{A}_k \bar{T}_k C^{\mathrm{T}} - \bar{\Theta}_k)(\bar{\Phi}_k - C\bar{T}_k C^{\mathrm{T}})^{+} \\
\bar{K} &= -(\bar{A}\bar{T}C^{\mathrm{T}} - \bar{\Theta})(\bar{\Phi} - C\bar{T}C^{\mathrm{T}})^{+}
\end{aligned}
$$

Moreover, we have two additional results:

Theorem 2.5. *The matrix $\bar{T}_k$ satisfies $\bar{T}_k = 0$ (and therefore $\bar{T} = 0$).*

Proof. Because $\bar{K}_k = 0$ (see Theorem 2.1), $\bar{T}_{k+1}$ can be expressed as

$$\bar{T}_{k+1} = \bar{A}_k \bar{T}_k \bar{A}_k^{\mathrm{T}} - \bar{K}_k(\bar{A}_k \bar{T}_k C^{\mathrm{T}} - \bar{\Theta}_k)^{\mathrm{T}} = \bar{A}_k \bar{T}_k \bar{A}_k^{\mathrm{T}}$$

The result then follows by induction, since $\bar{T}_0 = 0$. ∎

Theorem 2.6. *The matrix $\bar{\Theta}_k$ satisfies $\bar{\Theta}_k = 0$ (and therefore $\bar{\Theta} = 0$).*

Proof.

$$
\begin{aligned}
\bar{\Theta}_k &= \bar{A}_k \bar{P}_k C^{\mathrm{T}} + \bar{G}_k \bar{S} H^{\mathrm{T}} \\
&= A\Sigma_{k|k-1} C^{\mathrm{T}} + GSH^{\mathrm{T}} - K_k(C\Sigma_{k|k-1} C^{\mathrm{T}} + HRH^{\mathrm{T}}) \\
&= A\Sigma_{k|k-1} C^{\mathrm{T}} + GSH^{\mathrm{T}} \\
&\quad - (A\Sigma_{k|k-1} C^{\mathrm{T}} + GSH^{\mathrm{T}})(C\Sigma_{k|k-1} C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+}(C\Sigma_{k|k-1} C^{\mathrm{T}} + HRH^{\mathrm{T}}) \\
&= A\Sigma_{k|k-1} C^{\mathrm{T}} + GSH^{\mathrm{T}} - (A\Sigma_{k|k-1} C^{\mathrm{T}} + GSH^{\mathrm{T}}) = 0
\end{aligned}
$$

We have used (2.9) and the identity

$$
\begin{aligned}
(A\Sigma_{k|k-1} C^{\mathrm{T}} + GSH^{\mathrm{T}})(C\Sigma_{k|k-1} C^{\mathrm{T}} + HRH^{\mathrm{T}})^{+}(C\Sigma_{k|k-1} C^{\mathrm{T}} + HRH^{\mathrm{T}}) \\
= A\Sigma_{k|k-1} C^{\mathrm{T}} + GSH^{\mathrm{T}}
\end{aligned}
$$

which follows from Corollary 2.10 in Appendix 2.A, since

$$\begin{bmatrix} A\Sigma_{k|k-1}A^{\mathrm{T}} + GQG^{\mathrm{T}} & A\Sigma_{k|k-1}C^{\mathrm{T}} + GSH^{\mathrm{T}} \\ C\Sigma_{k|k-1}A^{\mathrm{T}} + HS^{\mathrm{T}}G^{\mathrm{T}} & C\Sigma_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}} \end{bmatrix}$$
$$= \begin{bmatrix} A & G & 0 \\ C & 0 & H \end{bmatrix} \begin{bmatrix} \Sigma_{k|k-1} & 0 & 0 \\ 0 & Q & S \\ 0 & S^{\mathrm{T}} & R \end{bmatrix} \begin{bmatrix} A^{\mathrm{T}} & C^{\mathrm{T}} \\ G^{\mathrm{T}} & 0 \\ 0 & H^{\mathrm{T}} \end{bmatrix} \geq 0 \qquad \blacksquare$$

2.3.3 Calculating $\mathbb{K}$ from (Φ, Θ)

The results of Section 2.3.1 extend to the $\mathcal{K}$-error system:

$$\begin{aligned}
\mathbb{T}_0 &= 0_{n\times n} \\
\mathbb{T}_{k+1} &= \mathbb{A}\mathbb{T}_k\mathbb{A}^{\mathrm{T}} + (\mathbb{A}\mathbb{T}_kC^{\mathrm{T}} - \Theta_k)(\Phi_k - C\mathbb{T}_kC^{\mathrm{T}})^{+}(\mathbb{A}\mathbb{T}_kC^{\mathrm{T}} - \Theta_k)^{\mathrm{T}} \qquad (2.37)\\
&= \mathrm{E}[(\hat{\mathrm{e}}_{k+1|k} - \mathbb{\mu}_{k+1})(\hat{\mathrm{e}}_{k+1|k} - \mathbb{\mu}_{k+1})^{\mathrm{T}}] \\
\Phi_k &\coloneqq C\mathbb{P}_kC^{\mathrm{T}} + HRH^{\mathrm{T}} \to \Phi \qquad (2.38)\\
\Theta_k &\coloneqq \mathbb{A}\mathbb{P}_kC^{\mathrm{T}} + \mathbb{G}\bar{S}H^{\mathrm{T}} \to \Theta \qquad (2.39)\\
\mathbb{T}_k &= \mathbb{P}_k - \mathbb{\Sigma}_{k|k-1} \to \mathbb{T} \coloneqq \mathbb{P} - \mathbb{\Sigma} = \mathbb{A}\mathbb{T}\mathbb{A}^{\mathrm{T}} - \mathbb{K}(\mathbb{A}\mathbb{T}C^{\mathrm{T}} - \Theta)^{\mathrm{T}} \qquad (2.40)\\
\mathbb{K}_k &= -(\mathbb{A}\mathbb{T}_kC^{\mathrm{T}} - \Theta_k)(\Phi_k - C\mathbb{T}_kC^{\mathrm{T}})^{+} \qquad (2.41)\\
\mathbb{K} &= -(\mathbb{A}\mathbb{T}C^{\mathrm{T}} - \Theta)(\Phi - C\mathbb{T}C^{\mathrm{T}})^{+}
\end{aligned}$$

2.3.4 Calculating $(\mathbb{K}, \Phi, \Theta, \mathbb{T})$ from (Φ, Θ)

Next we show how to calculate $\mathbb{K}_k$, Φ_k, Θ_k, and $\mathbb{T}_k$ from Φ_k and Θ_k. We also give the steady-state counterpart, showing how to calculate $\mathbb{K}$, Φ, Θ, and $\mathbb{T}$ from Φ and Θ.

Theorem 2.7. *The following equations hold:*

$$\begin{aligned}
\mathbb{K}_k &= K_k - \mathcal{K}(\Phi_k - CT_kC^{\mathrm{T}})(\Phi_k - CT_kC^{\mathrm{T}})^{+} \qquad (2.42)\\
\Phi_k &= \Phi_k + C(\mathbb{T}_k - T_k)C^{\mathrm{T}} \qquad (2.43)\\
\Theta_k &= \Theta_k + \mathbb{A}(\mathbb{T}_k - T_k)C^{\mathrm{T}} - \mathcal{K}\Phi_k \qquad (2.44)\\
\mathbb{T}_{k+1} &= \mathbb{A}\mathbb{T}_k\mathbb{A}^{\mathrm{T}} - \mathbb{K}_k(\mathbb{A}T_kC^{\mathrm{T}} - \Theta_k + \mathcal{K}\Phi_k)^{\mathrm{T}} \qquad (2.45)
\end{aligned}$$

If A is stable, then additionally

$$\mathbb{K} = K - \mathcal{K}(\Phi - CTC^{\mathrm{T}})(\Phi - CTC^{\mathrm{T}})^{+} \tag{2.46}$$

$$\Phi = \Phi + C(\mathbb{T} - T)C^{\mathrm{T}} \tag{2.47}$$

$$\Theta = \Theta + \mathbb{A}(\mathbb{T} - T)C^{\mathrm{T}} - \mathcal{K}\Phi \tag{2.48}$$

$$\mathbb{T} = \mathbb{A}\mathbb{T}\mathbb{A}^{\mathrm{T}} - \mathbb{K}(\mathbb{A}TC^{\mathrm{T}} - \Theta + \mathcal{K}\Phi)^{\mathrm{T}} \tag{2.49}$$

Proof. Using (2.13), (2.33), (2.35), (2.38), and (2.40), we obtain

$$\Phi_k - CT_kC^{\mathrm{T}} = C\Sigma_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}} = C\mathbb{\Sigma}_{k|k-1}C^{\mathrm{T}} + HRH^{\mathrm{T}} = \Phi_k - C\mathbb{T}_kC^{\mathrm{T}} \tag{2.50}$$

Combining (2.13), (2.35), and (2.40) gives

$$\mathbb{T}_k - T_k = \mathbb{P}_k - \mathbb{\Sigma}_{k|k-1} - (P_k - \Sigma_{k|k-1}) = \mathbb{P}_k - P_k \tag{2.51}$$

Using (2.50), we rewrite (2.15) as

$$\mathbb{K}_k = K_k - \mathcal{K}(\Phi_k - CT_kC^{\mathrm{T}})(\Phi_k - CT_kC^{\mathrm{T}})^{+}$$

Rearranging (2.50) yields

$$\Phi_k = \Phi_k + C(\mathbb{T}_k - T_k)C^{\mathrm{T}}$$

Rearranging (2.51), substituting into (2.39), and using (2.33) and (2.34) yields

$$\begin{aligned}
\Theta_k &= \mathbb{A}\mathbb{P}_kC^{\mathrm{T}} + GSH^{\mathrm{T}} - \mathcal{K}HRH^{\mathrm{T}} \\
&= \mathbb{A}(\mathbb{T}_k - T_k + P_k)C^{\mathrm{T}} + GSH^{\mathrm{T}} - \mathcal{K}HRH^{\mathrm{T}} \\
&= \mathbb{A}(\mathbb{T}_k - T_k)C^{\mathrm{T}} + \mathbb{A}P_kC^{\mathrm{T}} + GSH^{\mathrm{T}} - \mathcal{K}HRH^{\mathrm{T}} \\
&= \mathbb{A}(\mathbb{T}_k - T_k)C^{\mathrm{T}} + (A - \mathcal{K}C)P_kC^{\mathrm{T}} + GSH^{\mathrm{T}} - \mathcal{K}HRH^{\mathrm{T}} \\
&= AP_kC^{\mathrm{T}} + GSH^{\mathrm{T}} + \mathbb{A}(\mathbb{T}_k - T_k)C^{\mathrm{T}} - \mathcal{K}(CP_kC^{\mathrm{T}} + HRH^{\mathrm{T}}) \\
&= \Theta_k + \mathbb{A}(\mathbb{T}_k - T_k)C^{\mathrm{T}} - \mathcal{K}\Phi_k
\end{aligned}$$

Equations (2.43) and (2.44) depend on $\mathbb{T}_k$, which can be calculated from Φ_k and Θ_k by combining (2.37), (2.41), and (2.44):

$$\begin{aligned}
\mathbb{T}_{k+1} &= \mathbb{A}\mathbb{T}_k\mathbb{A}^{\mathrm{T}} - \mathbb{K}_k(\mathbb{A}\mathbb{T}_kC^{\mathrm{T}} - \mathbb{\Theta}_k)^{\mathrm{T}} \\
&= \mathbb{A}\mathbb{T}_k\mathbb{A}^{\mathrm{T}} - \mathbb{K}_k(\mathbb{A}\mathbb{T}_kC^{\mathrm{T}} - \Theta_k - \mathbb{A}(\mathbb{T}_k - T_k)C^{\mathrm{T}} + \mathcal{K}\Phi_k)^{\mathrm{T}} \\
&= \mathbb{A}\mathbb{T}_k\mathbb{A}^{\mathrm{T}} - \mathbb{K}_k(\mathbb{A}T_kC^{\mathrm{T}} - \Theta_k + \mathcal{K}\Phi_k)^{\mathrm{T}}
\end{aligned}$$

The steady-state versions, (2.46)–(2.49), are obtained by taking limits the limits of (2.42)–(2.45). Note that (2.46) relies on Assumption 3. Moreover, A must be stable for (2.57)–(2.59) to be relevant, because if A is not stable, then steady-state P, T, Φ, and Θ do not exist. If A is not stable but $\mathbb{A}$ is stable, then the quantities $\mathbb{K}$, Φ, Θ, and $\mathbb{T}$ exist, they just cannot be expressed as in (2.46)–(2.49). Equation (2.49) is a discrete Lyapunov equation with respect to $\mathbb{T}$. Because $\mathbb{A}$ is stable, it has a unique solution (see Corollary 2.9). ■

2.3.5 Calculating (K, Φ, Θ, T) from (Φ, Θ)

Similarly, K_k, Φ_k, Θ_k, and T_k can be calculated from Φ_k and Θ_k, and K, Φ, Θ, and T can be calculated from Φ and Θ:

Theorem 2.8. *The following equations hold:*

$$K_k = \mathbb{K}_k + \mathcal{K}(\Phi_k - C\mathbb{T}_k C^{\mathrm{T}})(\Phi_k - C\mathbb{T}_k C^{\mathrm{T}})^{+} \tag{2.52}$$

$$\Phi_k = \Phi_k + C(T_k - \mathbb{T}_k)C^{\mathrm{T}} \tag{2.53}$$

$$\Theta_k = \Theta_k + A(T_k - \mathbb{T}_k)C^{\mathrm{T}} + \mathcal{K}\Phi_k \tag{2.54}$$

$$T_{k+1} = AT_kA^{\mathrm{T}} - K_k(A\mathbb{T}_kC^{\mathrm{T}} - \Theta_k - \mathcal{K}\Phi_k)^{\mathrm{T}} \tag{2.55}$$

$$K = \mathbb{K} + \mathcal{K}(\Phi - C\mathbb{T}C^{\mathrm{T}})(\Phi - C\mathbb{T}C^{\mathrm{T}})^{+} \tag{2.56}$$

If A is stable, then additionally

$$\Phi = \Phi + C(T - \mathbb{T})C^{\mathrm{T}} \tag{2.57}$$

$$\Theta = \Theta + A(T - \mathbb{T})C^{\mathrm{T}} + \mathcal{K}\Phi \tag{2.58}$$

$$T = ATA^{\mathrm{T}} - K(A\mathbb{T}C^{\mathrm{T}} - \Theta - \mathcal{K}\Phi)^{\mathrm{T}} \tag{2.59}$$

Proof. First note

$$A = \mathbb{A} + \mathcal{K}C \tag{2.60}$$

$$GSH^{\mathrm{T}} = \mathbb{G}\bar{S}H^{\mathrm{T}} + \mathcal{K}HRH^{\mathrm{T}} \tag{2.61}$$

Rearranging (2.42) and using (2.50) yields

$$K_k = \mathbb{K}_k + \mathcal{K}(\Phi_k - C\mathbb{T}_k C^{\mathrm{T}})(\Phi_k - C\mathbb{T}_k C^{\mathrm{T}})^{+}$$

Rearranging (2.43) yields

$$\Phi_k = \Phi_k + C(T_k - \mathbb{T}_k)C^{\mathrm{T}}$$

Rearranging (2.51), substituting into (2.34), and using (2.38), (2.39), (2.60), and (2.61) yields

$$\begin{aligned}
\Theta_k &= AP_kC^{\mathrm{T}} + GSH^{\mathrm{T}} \\
&= A(\mathbb{P}_k + T_k - \mathbb{T}_k)C^{\mathrm{T}} + GSH^{\mathrm{T}} \\
&= A(T_k - \mathbb{T}_k)C^{\mathrm{T}} + A\mathbb{P}_kC^{\mathrm{T}} + GSH^{\mathrm{T}} \\
&= A(T_k - \mathbb{T}_k)C^{\mathrm{T}} + (\mathbb{A} + \mathcal{K}C)\mathbb{P}_kC^{\mathrm{T}} + \mathbb{G}\bar{S}H^{\mathrm{T}} + \mathcal{K}HRH^{\mathrm{T}} \\
&= \mathbb{A}\mathbb{P}_kC^{\mathrm{T}} + \mathbb{G}\bar{S}H^{\mathrm{T}} + A(T_k - \mathbb{T}_k)C^{\mathrm{T}} + \mathcal{K}(C\mathbb{P}_kC^{\mathrm{T}} + HRH^{\mathrm{T}}) \\
&= \mathbb{\Theta}_k + A(T_k - \mathbb{T}_k)C^{\mathrm{T}} + \mathcal{K}\mathbb{\Phi}_k
\end{aligned}$$

Equations (2.53) and (2.54) depend on T_k, which can be calculated from $\mathbb{\Phi}_k$ and $\mathbb{\Theta}_k$ by combining (2.32), (2.36), and (2.54):

$$\begin{aligned}
T_{k+1} &= AT_kA^{\mathrm{T}} - K_k(AT_kC^{\mathrm{T}} - \Theta_k)^{\mathrm{T}} \\
&= AT_kA^{\mathrm{T}} - K_k(AT_kC^{\mathrm{T}} - \mathbb{\Theta}_k - A(T_k - \mathbb{T}_k)C^{\mathrm{T}} - \mathcal{K}\mathbb{\Phi}_k)^{\mathrm{T}} \\
&= AT_kA^{\mathrm{T}} - K_k(A\mathbb{T}_kC^{\mathrm{T}} - \mathbb{\Theta}_k - \mathcal{K}\mathbb{\Phi}_k)^{\mathrm{T}}
\end{aligned}$$

The steady-state versions, (2.56)–(2.59), are obtained by taking the limits of (2.52)–(2.55). Note that (2.56) relies on Assumption 3. Moreover, A must be stable for (2.57)–(2.59) to be relevant, because if A is not stable, then steady-state P, T, Φ, and Θ do not exist. Equation (2.59) is a discrete Lyapunov equation with respect to T, so if A is stable, it has a unique solution (see Corollary 2.9). ■

2.4 Uniqueness of autocovariance parameterizations

In this section, we discuss the relationship between (Q, R, S) and (Φ, Θ) and, analogously, the relationship between (Q, R, S) and $(\mathbb{\Phi}, \mathbb{\Theta})$. We also discuss how assuming $S = 0$ affects these relationships.

2.4.1 Parameters (Q, R, S) and (Φ, Θ)

Both (Q, R, S) and (Φ, Θ) parameterize the steady-state output autocovariance displayed in (2.20)–(2.22). The matrices Q, R, and S are the fundamental stochastic parameters driving the system, and it is clear from (2.23) and (2.24) that (Φ, Θ)

is uniquely determined by (Q, R, S). However, it is not immediately clear if the converse is true. It proves useful to further characterize the relationship between (Q, R, S) and (Φ, Θ). To this end, we formally define the functions that map (Q, R, S) to Φ and Θ:

$$f_\Phi(Q, R, S) \coloneqq CPC^{\mathrm{T}} + HRH^{\mathrm{T}} \qquad f_\Theta(Q, R, S) \coloneqq APC^{\mathrm{T}} + GSH^{\mathrm{T}}$$

The domain of both f_Φ and f_Θ is taken to be the set of all (Q, R, S) where Q and R are symmetric, and P is understood to be the unique solution to (2.18). Both f_Φ and f_Θ are linear functions of (Q, R, S), which can be seen by vectorizing:

$$\begin{bmatrix} [f_\Phi(Q,R,S)]_{\mathrm{ss}} \\ [f_\Theta(Q,R,S)]_{\mathrm{s}} \end{bmatrix} = U \begin{bmatrix} Q_{\mathrm{ss}} \\ R_{\mathrm{ss}} \\ S_{\mathrm{s}} \end{bmatrix}$$

$$U \coloneqq \begin{bmatrix} U_{\Phi Q} & U_{\Phi R} & 0_{p(p+1)/2,gh} \\ U_{\Theta Q} & 0_{np,h(h+1)/2} & U_{\Theta S} \end{bmatrix} \tag{2.62}$$

$$U_{\Phi Q} \coloneqq \mathscr{D}_p^+(C \otimes C)F(G \otimes G)\mathscr{D}_g \quad U_{\Phi R} \coloneqq \mathscr{D}_p^+(H \otimes H)\mathscr{D}_h$$

$$U_{\Theta Q} \coloneqq (C \otimes A)F(G \otimes G)\mathscr{D}_g \qquad U_{\Theta S} \coloneqq H \otimes G$$

Next we examine the converse question: does (Φ, Θ) correspond to a unique value of (Q, R, S)? Furthermore, we have already shown that (Φ, Θ) uniquely determines K; we also investigate whether (Φ, Θ) uniquely determines L. To this end, we define the following set-valued functions, which return the sets of all values of (Q, R, S) and L corresponding to (Φ, Θ):

$$F_{QRS}(\Phi, \Theta) \coloneqq \{(Q, R, S) : f_\Phi(Q, R, S) = \Phi \text{ and } f_\Theta(Q, R, S) = \Theta\}$$

$$F_L(\Phi, \Theta) \coloneqq \{f_L(Q, R, S) : (Q, R, S) \in F_{QRS}(\Phi, \Theta) \text{ and } \begin{bmatrix} Q & S \\ S^{\mathrm{T}} & R \end{bmatrix} \geq 0\}$$

The domain of both F_{QRS} and F_L is taken to be the set of all (Φ, Θ) where Φ is symmetric. The members of the set $F_{QRS}(\Phi, \Theta)$ are the solutions to the linear system of equations

$$\begin{bmatrix} \Phi_{\mathrm{ss}} \\ \Theta_{\mathrm{s}} \end{bmatrix} = U \begin{bmatrix} Q_{\mathrm{ss}} \\ R_{\mathrm{ss}} \\ S_{\mathrm{s}} \end{bmatrix}$$

Thus, using Magnus and Neudecker (2019, Theorem 2.12), $F_{QRS}(\Phi, \Theta)$ can be written

as

$$F_{QRS}(\Phi,\Theta) = \{(Q,R,S) : \begin{bmatrix} Q_{\text{ss}} \\ R_{\text{ss}} \\ S_{\text{s}} \end{bmatrix} \in F^{\text{s}}_{QRS}(\Phi,\Theta)\} \tag{2.63}$$

$$F^{\text{s}}_{QRS}(\Phi,\Theta) := \begin{cases} F^{\star}_{QRS}(\Phi,\Theta), & UU^{+}\begin{bmatrix} \Phi_{\text{ss}} \\ \Theta_{\text{s}} \end{bmatrix} = \begin{bmatrix} \Phi_{\text{ss}} \\ \Theta_{\text{s}} \end{bmatrix} \\ \emptyset, & \text{otherwise} \end{cases} \tag{2.64}$$

$$F^{\star}_{QRS}(\Phi,\Theta) := \{U^{+}\begin{bmatrix} \Phi_{\text{ss}} \\ \Theta_{\text{s}} \end{bmatrix} + (I - U^{+}U)q : q \text{ arbitrary}\} \tag{2.65}$$

The set $F_{QRS}(\Phi,\Theta)$ has zero, one, or infinitely many members, depending on the linear independence of the columns of U and the value of (Φ,Θ). A caveat is that if U has LI rows, then $UU^{+} = I$, thus $F_{QRS}(\Phi,\Theta) \neq \emptyset$ and the value of (Φ,Θ) has no bearing on the cardinality of $F_{QRS}(\Phi,\Theta)$.

Remark 2.3. There exist systems with all possible combinations of linear independence/dependence of the rows and columns of U in (2.62). Examples are given below:

A	C	G	H	U rows	U columns
$\begin{bmatrix} 0.5 & 0 \\ 1 & 0.5 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$	1	LI	LI
0.5	1	1	1	LI	LD
$\begin{bmatrix} 0.5 & 0 \\ 1 & 0.5 \end{bmatrix}$	I_2	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$	LD	LI
$\begin{bmatrix} 0.5 & 0 \\ 1 & 0.5 \end{bmatrix}$	I_2	I_2	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$	LD	LD

Next, we give an alternate characterization of $F_{QRS}(\Phi,\Theta)$:

Theorem 2.9. *For any (Φ,Θ) with Φ symmetric, $(Q,R,S) \in F_{QRS}(\Phi,\Theta)$ if and only if there exists a symmetric matrix P such that Q, R, S, and P satisfy*

$$P = APA^{\text{T}} + GQG^{\text{T}} \tag{2.18}$$

$$\Phi = CPC^{\text{T}} + HRH^{\text{T}} \tag{2.23}$$

$$\Theta = APC^{\text{T}} + GSH^{\text{T}} \tag{2.24}$$

Proof. First, $f_\Phi(Q,R,S) = \Phi$ and $f_\Theta(Q,R,S) = \Theta$ are precisely (2.23) and (2.24), respectively, with (2.18) holding implicitly. Because of the domain of both f_Φ and f_Θ, GQG^{T} is symmetric, which implies that P is symmetric (see Corollary 2.9). Conversely, if (2.18), (2.23), and (2.24) hold for some symmetric P, then clearly $f_\Phi(Q,R,S) = \Phi$ and $f_\Theta(Q,R,S) = \Theta$. ■

Corollary 2.3. *Assume $G = I_n$ and $H = I_p$. Then (Φ, Θ) does not uniquely determine (Q,R,S). Specifically, $F_{QRS}(\Phi,\Theta)$ is an infinite set, and, equivalently, U has LI rows and LD columns.*

Proof. Let P be an arbitrary symmetric matrix, and let Q, R, and S be

$$Q = P - APA^{\mathrm{T}} \qquad R = \Phi - CPC^{\mathrm{T}} \qquad S = \Theta - APC^{\mathrm{T}}$$

Then $f_\Phi(Q,R,S) = \Phi$ and $f_\Theta(Q,R,S) = \Theta$. Because $G = I_n$, $P \leftrightarrow Q$ is bijective (see (2.19)). Thus arbitrary symmetric P corresponds to infinitely many values of (Q,R,S). ■

Remark 2.4. Assume $G = I_n$ and $H = I_p$. Then, in general, (Φ,Θ) does not uniquely determine L, i.e., the set $F_L(\Phi,\Theta)$ has more than one member. To illustrate this point, consider the system

$$A = \begin{bmatrix} 0.5 & 0 \\ 1 & 0.5 \end{bmatrix} \qquad C = \begin{bmatrix} 1 & 1 \end{bmatrix} \qquad G = I_2 \qquad H = 1$$

with two sets of covariance matrices:

$$Q^\star = I_2, \quad R^\star = 1, \quad S^\star = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\tilde{Q} = \begin{bmatrix} 0.3 & 0.1 \\ 0.1 & 2.2 \end{bmatrix}, \quad \tilde{R} = 3.2074, \quad \tilde{S} = \begin{bmatrix} 0.7111 \\ 1.8148 \end{bmatrix}$$

For both sets, Φ, Θ, and K are the same, but the L's are different:

$$f_\Phi(Q^\star,R^\star,S^\star) = f_\Phi(\tilde{Q},\tilde{R},\tilde{S}) = 8.4074$$

$$f_\Theta(Q^\star,R^\star,S^\star) = f_\Theta(\tilde{Q},\tilde{R},\tilde{S}) = \begin{bmatrix} 1.1111 \\ 4.8148 \end{bmatrix}$$

$$f_K(Q^\star,R^\star,S^\star) = f_K(\tilde{Q},\tilde{R},\tilde{S}) = \begin{bmatrix} 0.1694 \\ 0.5508 \end{bmatrix}$$

$$f_L(Q^\star,R^\star,S^\star) = \begin{bmatrix} 0.3388 \\ 0.4241 \end{bmatrix} \neq f_L(\tilde{Q},\tilde{R},\tilde{S}) = \begin{bmatrix} 0.0015 \\ 0.2380 \end{bmatrix}$$

Remark 2.5. In (2.63)–(2.65) and Theorem 2.9, we have given two explicit characterizations of the solution to the unconstrained version of the stochastic realization problem from Faurre (1976). In this paper, Faurre discusses the constrained stochastic realization problem, which is solved by

$$F^{\mathrm{c}}_{QRS}(\Phi,\Theta) = \{(Q,R,S) : (Q,R,S) \in F_{QRS}(\Phi,\Theta),\ \begin{bmatrix} Q & S \\ S^{\mathrm{T}} & R \end{bmatrix} \geq 0,\ P > 0\}$$

The matrices (Q,R,S) form a valid joint covariance matrix if and only if the first semidefinite constraint is satisfied. The second constraint, $P > 0$, is an artifact of Faurre's work in the sense that he imposes this constraint because the algorithms which he provides to solve the constrained stochastic realization problem require P to be invertible. In practice, however, it is possible that P may be singular, but positive semidefinite, i.e., $P \geq 0$. One of Faurre's contributions is that if $G = I_n$ and $H = I_p$, then $F^{\mathrm{c}}_{QRS}(\Phi,\Theta)$ is a closed and bounded set. In contrast, $F_{QRS}(\Phi,\Theta)$ is often unbounded. Since covariance estimation is our primary concern in this dissertation, we apply semidefinite constraints only to the covariance estimation problem, and not to the stochastic realization problem. For example, the constrained ALS problem is presented in Section 3.4. It is the nature of the linear relationship between (Q,R,S) and (Φ,Θ) that affects the positive definiteness of the Hessian (i.e., the uniqueness) for the ALS problem, not the semidefinite constraints.

2.4.2 Parameters $(Q,R,0)$ and (Φ,Θ)

Oftentimes S is assumed to be a zero matrix. Next, we investigate whether any of the results of Section 2.4.1 change in this case. Define the functions that map $(Q,R,0)$ to Φ and Θ:

$$\tilde{f}_\Phi(Q,R) := f_\Phi(Q,R,0) \qquad \tilde{f}_\Theta(Q,R) := f_\Theta(Q,R,0)$$

The domain of both $\tilde{f}_\Phi$ and $\tilde{f}_\Theta$ is taken to be the set of all (Q,R) where Q and R are symmetric, and P is the unique solution to (2.18). The vectorized versions of $\tilde{f}_\Phi$ and $\tilde{f}_\Theta$ are

$$\begin{bmatrix} [\tilde{f}_\Phi(Q,R)]_{\mathrm{ss}} \\ [\tilde{f}_\Theta(Q,R)]_{\mathrm{s}} \end{bmatrix} = \tilde{U} \begin{bmatrix} Q_{\mathrm{ss}} \\ R_{\mathrm{ss}} \end{bmatrix}$$
$$\tilde{U} := \begin{bmatrix} U_{\Phi Q} & U_{\Phi R} \\ U_{\Theta Q} & 0_{np,h(h+1)/2} \end{bmatrix} \tag{2.66}$$

We define the following set-valued functions, which return the sets of all values of (Q, R) and L corresponding to (Φ, Θ) when $S = 0$:

$$F_{QR}(\Phi, \Theta) \coloneqq \{(Q, R) : \tilde{f}_\Phi(Q, R) = \Phi \text{ and } \tilde{f}_\Theta(Q, R) = \Theta\}$$
$$\tilde{F}_L(\Phi, \Theta) \coloneqq \{\tilde{f}_L(Q, R) : (Q, R) \in F_{QR}(\Phi, \Theta) \text{ and } Q \geq 0,\ R \geq 0\}$$

The domain of both F_{QR} and $\tilde{F}_L$ is taken to be the set of all (Φ, Θ) where Φ is symmetric. The members of the set $F_{QR}(\Phi, \Theta)$ are the solutions to the linear system of equations

$$\begin{bmatrix} \Phi_{\text{ss}} \\ \Theta_{\text{s}} \end{bmatrix} = \tilde{U} \begin{bmatrix} Q_{\text{ss}} \\ R_{\text{ss}} \end{bmatrix}$$

Therefore $F_{QR}(\Phi, \Theta)$ can be written as

$$F_{QR}(\Phi, \Theta) = \{(Q, R) : \begin{bmatrix} Q_{\text{ss}} \\ R_{\text{ss}} \end{bmatrix} \in F^{\text{s}}_{QR}(\Phi, \Theta)\}$$

$$F^{\text{s}}_{QR}(\Phi, \Theta) \coloneqq \begin{cases} F^{\star}_{QR}(\Phi, \Theta), & \tilde{U}\tilde{U}^{+} \begin{bmatrix} \Phi_{\text{ss}} \\ \Theta_{\text{s}} \end{bmatrix} = \begin{bmatrix} \Phi_{\text{ss}} \\ \Theta_{\text{s}} \end{bmatrix} \\ \emptyset, & \text{otherwise} \end{cases}$$

$$F^{\star}_{QR}(\Phi, \Theta) \coloneqq \{\tilde{U}^{+} \begin{bmatrix} \Phi_{\text{ss}} \\ \Theta_{\text{s}} \end{bmatrix} + (I - \tilde{U}^{+}\tilde{U})q : q \text{ arbitrary}\}$$

As with $F_{QRS}(\Phi, \Theta)$, the set $F_{QR}(\Phi, \Theta)$ has zero, one, or infinitely many members, depending on the linear independence of the columns of $\tilde{U}$ and the value of (Φ, Θ). If $\tilde{U}$ has LI rows, then $\tilde{U}\tilde{U}^{+} = I$, thus $F_{QR}(\Phi, \Theta) \neq \emptyset$ and the value of (Φ, Θ) has no bearing on the cardinality of $F_{QR}(\Phi, \Theta)$.

Remark 2.6. There exist systems with all possible combinations of linear independence/dependence of the rows and columns of $\tilde{U}$ in (2.66). Examples for each case are given below:

A	C	G	H	$\tilde{U}$ rows	$\tilde{U}$ columns
0.5	1	1	1	LI	LI
$\begin{bmatrix} 0.5 & 0 \\ 1 & 0.5 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 \end{bmatrix}$	I_2	1	LI	LD
$\begin{bmatrix} 0.5 & 0 \\ 1 & 0.5 \end{bmatrix}$	I_2	I_2	I_2	LD	LI
$\begin{bmatrix} 0.5 & 0 & 0 \\ 1 & 0.5 & 0 \\ 1 & 0 & 0.5 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$	I_3	I_2	LD	LD

In Section 2.4.1, we treated the special case of $G = I_n$ and $H = I_p$ to show these conditions imply U has LI rows and LD columns. Clearly this result does not also apply to $\tilde{U}$. The examples in Remark 2.6 all have $G = I_n$ and $H = I_p$, but different combinations of the linear dependence for the rows and columns of $\tilde{U}$.

Next, we provide counterparts to Theorem 2.9, Corollary 2.3, and Remark 2.4.

Theorem 2.10. *For any (Φ, Θ) with Φ symmetric, $(Q, R) \in F_{QR}(\Phi, \Theta)$ if and only if there exists a symmetric matrix P such that Q, R, and P satisfy*

$$P = APA^{\mathrm{T}} + GQG^{\mathrm{T}} \qquad \Phi = CPC^{\mathrm{T}} + HRH^{\mathrm{T}} \qquad \Theta = APC^{\mathrm{T}}$$

Proof of Theorem 2.10 is analogous to that for Theorem 2.9.

Corollary 2.4. *Assume $G = I_n$, $H = I_p$, and $S = 0$. Then (Φ, Θ) uniquely determines (Q, R) ($\tilde{U}$ has LI columns) if and only if A is invertible and C has LI columns.*

Proof. We proceed by characterizing the null space of $\tilde{U}$. Consider

$$0 = \tilde{U} \begin{bmatrix} Q_{\mathrm{ss}} \\ R_{\mathrm{ss}} \end{bmatrix}$$

the unvectorized form of which may be written as three separate equations:

$$P = APA^{\mathrm{T}} + Q \quad (2.67) \qquad 0 = CPC^{\mathrm{T}} + R \quad (2.68) \qquad 0 = APC^{\mathrm{T}} \quad (2.69)$$

Assume that A is invertible and C has LI columns. Then (2.69) implies $P = 0$, and then (2.67) and (2.68) imply $Q = 0$ and $R = 0$, respectively.

Assume A is not invertible. Let $z \neq 0$ such that $Az = 0$. Set $P = zz^{\mathrm{T}} \neq 0$, $Q = P$, and $R = -CPC^{\mathrm{T}}$. Then (2.67)–(2.69) are satisfied. Thus $\left[\begin{smallmatrix} Q_{\mathrm{ss}} \\ R_{\mathrm{ss}} \end{smallmatrix}\right]$ is a nonzero element of null $\tilde{U}$.

Assume C has LD columns. Let $z \neq 0$ such that $Cz = 0$. Set $P = zz^{\mathrm{T}} \neq 0$, $Q = P - APA^{\mathrm{T}}$, and $R = 0$. Then (2.67)–(2.69) are satisfied. Moreover, $P \neq 0 \implies Q \neq 0$ (recall (2.19)). Thus $\left[\begin{smallmatrix} Q_{\mathrm{ss}} \\ R_{\mathrm{ss}} \end{smallmatrix}\right]$ is a nonzero element of null $\tilde{U}$. ■

Remark 2.7. If $S = 0$ and A is invertible, then (Φ, Θ) uniquely determines L. Observe that $S = 0$ implies $AL = K$, and then A invertible implies $L = A^{-1}K$, and (Φ, Θ) uniquely determines K (recall Theorem 2.3). Furthermore, if additionally $H = I_p$, then (Φ, Θ) also uniquely determines R. Observe that $S = 0$ and A invertible together imply $PC^{\mathrm{T}} = A^{-1}\Theta$. Then $H = I_p$ implies $R = \Phi - CPC^{\mathrm{T}} = \Phi - CA^{-1}\Theta$. To illustrate these points, consider the system from Remark 2.4 with a third set of covariance matrices:

$$Q^{\#} = \begin{bmatrix} 2.5 & -2.5 \\ -2.5 & 2.5 \end{bmatrix}, \quad R^{\#} = 1, \quad S^{\#} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Then $Q^{\#} \neq Q^{\star}$, but $R^{\#} = R^{\star}$, $\tilde{f}_{\Phi}(Q^{\#}, R^{\#}) = \tilde{f}_{\Phi}(Q^{\star}, R^{\star})$, $\tilde{f}_{\Theta}(Q^{\#}, R^{\#}) = \tilde{f}_{\Theta}(Q^{\star}, R^{\star})$, $\tilde{f}_{K}(Q^{\#}, R^{\#}) = \tilde{f}_{K}(Q^{\star}, R^{\star})$, and $\tilde{f}_{L}(Q^{\#}, R^{\#}) = \tilde{f}_{L}(Q^{\star}, R^{\star})$.

2.4.3 Parameters (Q, R, S) and (Φ, Θ)

The relationship between (Q, R, S) and (Φ, Θ) has analogous results to Remarks 2.3 and 2.4, Theorem 2.9, and Corollary 2.3. We give these results subsequently. Define the functions that map (Q, R, S) to Φ and Θ:

$$f_{\Phi}(Q, R, S) \coloneqq C\mathbb{P}C^{\mathrm{T}} + HRH^{\mathrm{T}} \qquad f_{\Theta}(Q, R, S) \coloneqq \mathbb{A}\mathbb{P}C^{\mathrm{T}} + GSH^{\mathrm{T}} - \mathcal{K}HRH^{\mathrm{T}}$$

The domain of both f_{Φ} and f_{Θ} is taken to be the set of all (Q, R, S) where Q and R are symmetric and $\mathbb{P}$ is the unique solution to (2.26). Both f_{Φ} and f_{Θ} are linear functions of (Q, R, S) which can be seen by vectorizing:

$$\begin{bmatrix} [f_{\Phi}(Q, R, S)]_{\mathrm{ss}} \\ [f_{\Theta}(Q, R, S)]_{\mathrm{s}} \end{bmatrix} = \mathbb{U} \begin{bmatrix} Q_{\mathrm{ss}} \\ R_{\mathrm{ss}} \\ S_{\mathrm{s}} \end{bmatrix}$$
$$\mathbb{U} \coloneqq \begin{bmatrix} \mathbb{U}_{\Phi Q} & \mathbb{U}_{\Phi R} & \mathbb{U}_{\Phi S} \\ \mathbb{U}_{\Theta Q} & \mathbb{U}_{\Theta R} & \mathbb{U}_{\Theta S} \end{bmatrix} \tag{2.70}$$

$$
\begin{aligned}
\mathbb{U}_{\Phi Q} &:= \mathscr{D}_p^+(C \otimes C)\mathbb{F}(G \otimes G)\mathscr{D}_g \\
\mathbb{U}_{\Phi R} &:= \mathscr{D}_p^+[(C \otimes C)\mathbb{F}(\mathcal{K}H \otimes \mathcal{K}H) + (H \otimes H)]\mathscr{D}_h \\
\mathbb{U}_{\Phi S} &:= -\mathscr{D}_p^+(C \otimes C)\mathbb{F}(I_{n^2} + \mathscr{K}_n)(\mathcal{K}H \otimes G) \\
\mathbb{U}_{\Theta Q} &:= (C \otimes \mathbb{A})\mathbb{F}(G \otimes G)\mathscr{D}_g \\
\mathbb{U}_{\Theta R} &:= [(C \otimes \mathbb{A})\mathbb{F}(\mathcal{K}H \otimes \mathcal{K}H) - (H \otimes \mathcal{K}H)]\mathscr{D}_h \\
\mathbb{U}_{\Theta S} &:= -(C \otimes \mathbb{A})\mathbb{F}(I_{n^2} + \mathscr{K}_n)(\mathcal{K}H \otimes G) + (H \otimes G)
\end{aligned}
$$

Define the following set-valued functions, which return the sets of all values of (Q, R, S) and L corresponding to (Φ, Θ):

$$
\begin{aligned}
\mathbb{F}_{QRS}(\Phi, \Theta) &:= \{(Q, R, S) : f_\Phi(Q, R, S) = \Phi \text{ and } f_\Theta(Q, R, S) = \Theta\} \\
\mathbb{F}_{L}(\Phi, \Theta) &:= \{f_L(Q, R, S) : (Q, R, S) \in \mathbb{F}_{QRS}(\Phi, \Theta) \text{ and } \left[\begin{smallmatrix} Q & S \\ S^{\mathrm{T}} & R \end{smallmatrix}\right] \geq 0\}
\end{aligned}
$$

The domain of both $\mathbb{F}_{QRS}$ and $\mathbb{F}_L$ is taken to be the set of all (Φ, Θ) where Φ is symmetric. The members of the set $\mathbb{F}_{QRS}(\Phi, \Theta)$ are the solutions to the linear system of equations

$$
\begin{bmatrix} \Phi_{\mathrm{ss}} \\ \Theta_{\mathrm{s}} \end{bmatrix} = \mathbb{U} \begin{bmatrix} Q_{\mathrm{ss}} \\ R_{\mathrm{ss}} \\ S_{\mathrm{s}} \end{bmatrix}
$$

Therefore $\mathbb{F}_{QRS}(\Phi, \Theta)$ can be written as

$$
\mathbb{F}_{QRS}(\Phi, \Theta) = \{(Q, R, S) : \begin{bmatrix} Q_{\mathrm{ss}} \\ R_{\mathrm{ss}} \\ S_{\mathrm{s}} \end{bmatrix} \in \mathbb{F}^{\mathrm{s}}_{QRS}(\Phi, \Theta)\} \tag{2.71}
$$

$$
\mathbb{F}^{\mathrm{s}}_{QRS}(\Phi, \Theta) := \begin{cases} \mathbb{F}^{\star}_{QRS}(\Phi, \Theta), & \mathbb{U}\mathbb{U}^{+} \begin{bmatrix} \Phi_{\mathrm{ss}} \\ \Theta_{\mathrm{s}} \end{bmatrix} = \begin{bmatrix} \Phi_{\mathrm{ss}} \\ \Theta_{\mathrm{s}} \end{bmatrix} \\ \emptyset, & \text{otherwise} \end{cases} \tag{2.72}
$$

$$
\mathbb{F}^{\star}_{QRS}(\Phi, \Theta) := \{\mathbb{U}^{+} \begin{bmatrix} \Phi_{\mathrm{ss}} \\ \Theta_{\mathrm{s}} \end{bmatrix} + (I - \mathbb{U}^{+}\mathbb{U})q : q \text{ arbitrary}\} \tag{2.73}
$$

The set $\mathbb{F}_{QRS}(\Phi, \Theta)$ has zero, one, or infinitely many members, depending on the linear independence of the columns of $\mathbb{U}$ and the value of (Φ, Θ). If $\mathbb{U}$ has LI rows, then $\mathbb{U}\mathbb{U}^{+} = I$, thus $\mathbb{F}_{QRS}(\Phi, \Theta) \neq \emptyset$ and the value of (Φ, Θ) has no bearing on the cardinality of $\mathbb{F}_{QRS}(\Phi, \Theta)$.

Remark 2.8. There exist systems with all possible combinations of linear independence/dependence of the rows and columns of $\mathbb{U}$ in (2.70). Moreover, if $\mathcal{K} = 0_{n,p}$, then $\mathbb{U} = U$. Therefore, the example systems from Remark 2.3 can be used again if one takes $\mathcal{K} = 0_{n,p}$. This choice for $\mathcal{K}$ results in a stable $\mathbb{A}$ for each example, since A is stable in each example.

Theorem 2.11. *For any* (Φ, Θ) *with* Φ *symmetric,* $(Q, R, S) \in \mathbb{F}_{QRS}(\Phi, \Theta)$ *if and only if there exists a symmetric matrix* $\mathbb{P}$ *such that* Q*,* R*,* S*, and* $\mathbb{P}$ *satisfy*

$$\mathbb{P} = \mathbb{A}\mathbb{P}\mathbb{A}^{\mathrm{T}} + GQG^{\mathrm{T}} + \mathcal{K}HRH^{\mathrm{T}}\mathcal{K}^{\mathrm{T}} - GSH^{\mathrm{T}}\mathcal{K}^{\mathrm{T}} - \mathcal{K}HS^{\mathrm{T}}G^{\mathrm{T}} \tag{2.26}$$

$$\Phi = C\mathbb{P}C^{\mathrm{T}} + HRH^{\mathrm{T}} \tag{2.30}$$

$$\Theta = A\mathbb{P}C^{\mathrm{T}} + GSH^{\mathrm{T}} - \mathcal{K}(C\mathbb{P}C^{\mathrm{T}} + HRH^{\mathrm{T}}) \tag{2.31}$$

Proof of Theorem 2.11 analogous to that for Theorem 2.9.

Corollary 2.5. *Assume* $G = I_n$ *and* $H = I_p$*. Then* (Φ, Θ) *does not uniquely determine* (Q, R, S)*. Specifically,* $\mathbb{F}_{QRS}(\Phi, \Theta)$ *is an infinite set, and, equivalently,* $\mathbb{U}$ *has LI rows and LD columns.*

Proof. Let $\mathbb{P}$ be an arbitrary symmetric matrix, and let Q, R, and S be

$$Q = \mathbb{P} - \mathbb{A}\mathbb{P}\mathbb{A}^{\mathrm{T}} - \mathcal{K}R\mathcal{K}^{\mathrm{T}} + S\mathcal{K}^{\mathrm{T}} + \mathcal{K}S^{\mathrm{T}}$$

$$R = \Phi - C\mathbb{P}C^{\mathrm{T}} \qquad S = \Theta - \mathbb{A}\mathbb{P}C^{\mathrm{T}} + \mathcal{K}R$$

Then $f_\Phi(Q, R, S) = \Phi$ and $f_\Theta(Q, R, S) = \Theta$. Because $\mathbb{P} \leftrightarrow Q + \mathcal{K}R\mathcal{K}^{\mathrm{T}} - S\mathcal{K}^{\mathrm{T}} - \mathcal{K}S^{\mathrm{T}}$ is bijective (see (2.27)), arbitrary symmetric $\mathbb{P}$ corresponds to infinitely many values of (Q, R, S), as it is impossible to form arbitrary symmetric values of $Q + \mathcal{K}R\mathcal{K}^{\mathrm{T}} - S\mathcal{K}^{\mathrm{T}} - \mathcal{K}S^{\mathrm{T}}$ from finitely many values of (Q, R, S). ∎

Remark 2.9. Assume $G = I_n$ and $H = I_p$. Then, in general, (Φ, Θ) does not uniquely determine L, i.e., the set $\mathbb{F}_L(\Phi, \Theta)$ has more than one member. The example system from Remark 2.4 can be used again if one takes $\mathcal{K} = 0_{n,p}$.

2.4.4 Parameters $(Q, R, 0)$ and (Φ, Θ)

With $S = 0$, the relationship between $(Q, R, 0)$ and (Φ, Θ) has analogous results to Remarks 2.6 and 2.7, Theorem 2.10, and Corollary 2.4. We give these results subsequently. Define the functions that map $(Q, R, 0)$ to Φ and Θ:

$$\tilde{f}_\Phi(Q, R) := f_\Phi(Q, R, 0) \qquad \tilde{f}_\Theta(Q, R) := f_\Theta(Q, R, 0)$$

The domain of both $\tilde{f}_\Phi$ and $\tilde{f}_\Theta$ is taken to be the set of all (Q, R) where Q and R are symmetric, and $\mathbb{P}$ is the unique solution to (2.26). The vectorized versions of $\tilde{f}_\Phi$ and $\tilde{f}_\Theta$ are

$$\begin{bmatrix} [\tilde{f}_\Phi(Q,R)]_{\text{ss}} \\ [\tilde{f}_\Theta(Q,R)]_{\text{s}} \end{bmatrix} = \tilde{\mathbb{U}} \begin{bmatrix} Q_{\text{ss}} \\ R_{\text{ss}} \end{bmatrix}$$
$$\tilde{\mathbb{U}} := \begin{bmatrix} \mathbb{U}_{\Phi Q} & \mathbb{U}_{\Phi R} \\ \mathbb{U}_{\Theta Q} & \mathbb{U}_{\Theta R} \end{bmatrix} \tag{2.74}$$

We define the following set-valued functions, which return the sets of all values of (Q, R) and L corresponding to (Φ, Θ) when $S = 0$:

$$\mathbb{F}_{QR}(\Phi, \Theta) := \{(Q,R) : \tilde{f}_\Phi(Q,R) = \Phi \text{ and } \tilde{f}_\Theta(Q,R) = \Theta\}$$
$$\tilde{\mathbb{F}}_L(\Phi, \Theta) := \{\tilde{f}_L(Q,R) : (Q,R) \in \mathbb{F}_{QR}(\Phi,\Theta) \text{ and } Q \geq 0,\ R \geq 0\}$$

The domain of both $\mathbb{F}_{QR}$ and $\tilde{\mathbb{F}}_L$ is taken to be the set of all (Φ, Θ) where Φ is symmetric. The members of the set $\mathbb{F}_{QR}(\Phi, \Theta)$ are the solutions to the linear system of equations

$$\begin{bmatrix} \Phi_{\text{ss}} \\ \Theta_{\text{s}} \end{bmatrix} = \tilde{\mathbb{U}} \begin{bmatrix} Q_{\text{ss}} \\ R_{\text{ss}} \end{bmatrix}$$

Therefore $\mathbb{F}_{QR}(\Phi, \Theta)$ can be written as

$$\mathbb{F}_{QR}(\Phi,\Theta) = \{(Q,R) : \begin{bmatrix} Q_{\text{ss}} \\ R_{\text{ss}} \end{bmatrix} \in \mathbb{F}^{\text{s}}_{QR}(\Phi,\Theta)\}$$
$$\mathbb{F}^{\text{s}}_{QR}(\Phi,\Theta) := \begin{cases} \mathbb{F}^{\star}_{QR}(\Phi,\Theta), & \tilde{\mathbb{U}}\tilde{\mathbb{U}}^{+} \begin{bmatrix} \Phi_{\text{ss}} \\ \Theta_{\text{s}} \end{bmatrix} = \begin{bmatrix} \Phi_{\text{ss}} \\ \Theta_{\text{s}} \end{bmatrix} \\ \emptyset, & \text{otherwise} \end{cases}$$
$$\mathbb{F}^{\star}_{QR}(\Phi,\Theta) := \{\tilde{\mathbb{U}}^{+} \begin{bmatrix} \Phi_{\text{ss}} \\ \Theta_{\text{s}} \end{bmatrix} + (I - \tilde{\mathbb{U}}^{+}\tilde{\mathbb{U}})q : q \text{ arbitrary}\}$$

As with $\mathbb{F}_{QRS}(\Phi, \Theta)$, the set $\mathbb{F}_{QR}(\Phi, \Theta)$ has zero, one, or infinitely many members, depending on the linear independence of the columns of $\tilde{\mathbb{U}}$ and the value of (Φ, Θ). If $\tilde{\mathbb{U}}$ has LI rows, then $\tilde{\mathbb{U}}\tilde{\mathbb{U}}^{+} = I$, thus $\mathbb{F}_{QR}(\Phi, \Theta) \neq \emptyset$ and the value of (Φ, Θ) has no bearing on the cardinality of $\mathbb{F}_{QR}(\Phi, \Theta)$.

Remark 2.10. There exist systems with all possible combinations of linear independence/dependence of the rows and columns of $\tilde{\mathbb{U}}$ in (2.74). Moreover, if $\mathcal{K} = 0_{n,p}$, then $\tilde{\mathbb{U}} = \tilde{U}$. Therefore, the example systems from Remark 2.6 can be used again if one takes $\mathcal{K} = 0_{n,p}$. This choice for $\mathcal{K}$ results in a stable $\mathbb{A}$ for each example, since A is stable in each example.

In Section 2.4.3, we treated the special case of $G = I_n$ and $H = I_p$ to show these conditions imply $\mathbb{U}$ has LI rows and LD columns. This result does not also apply to $\tilde{\mathbb{U}}$. The examples in Remark 2.10 all have $G = I_n$ and $H = I_p$, but different combinations of linear dependence for the rows and columns of $\tilde{\mathbb{U}}$.

Theorem 2.12. *For any (Φ, Θ) with Φ symmetric, $(Q, R) \in \mathbb{F}_{QR}(\Phi, \Theta)$ if and only if there exists a symmetric $\mathbb{P}$ such that Q, R, and $\mathbb{P}$ satisfy*

$$\mathbb{P} = \mathbb{A}\mathbb{P}\mathbb{A}^{\mathrm{T}} + GQG^{\mathrm{T}} + \mathcal{K}HRH^{\mathrm{T}}\mathcal{K}^{\mathrm{T}}$$
$$\Phi = C\mathbb{P}C^{\mathrm{T}} + HRH^{\mathrm{T}} \qquad \Theta = \mathbb{A}\mathbb{P}C^{\mathrm{T}} - \mathcal{K}HRH^{\mathrm{T}}$$

Proof of Theorem 2.12 is analogous to that for Theorem 2.11.

Corollary 2.6. *Assume $G = I_n$, $H = I_p$, and $S = 0$. Then (Φ, Θ) uniquely determines (Q, R) (i.e., $\tilde{\mathbb{U}}$ has LI columns) if and only if A is invertible and C has LI columns.*

Proof. We proceed by characterizing the null space of $\tilde{\mathbb{U}}$. Consider

$$0 = \tilde{\mathbb{U}} \begin{bmatrix} Q_{\mathrm{ss}} \\ R_{\mathrm{ss}} \end{bmatrix}$$

the unvectorized form of which may be written as three separate equations:

$$\begin{aligned}
\mathbb{P} &= \mathbb{A}\mathbb{P}\mathbb{A}^{\mathrm{T}} + Q + \mathcal{K}R\mathcal{K}^{\mathrm{T}} \\
&= A\mathbb{P}A^{\mathrm{T}} + \mathcal{K}(C\mathbb{P}C^{\mathrm{T}} + R)\mathcal{K}^{\mathrm{T}} + Q - \mathcal{K}C\mathbb{P}A^{\mathrm{T}} - A\mathbb{P}C^{\mathrm{T}}\mathcal{K}^{\mathrm{T}} && (2.75) \\
0 &= C\mathbb{P}C^{\mathrm{T}} + R && (2.76) \\
0 &= \mathbb{A}\mathbb{P}C^{\mathrm{T}} - \mathcal{K}R \\
&= A\mathbb{P}C^{\mathrm{T}} - \mathcal{K}(C\mathbb{P}C^{\mathrm{T}} + R) && (2.77)
\end{aligned}$$

Substituting (2.76) into (2.75) and (2.77) yields additionally

$$\begin{aligned}
\mathbb{P} &= A\mathbb{P}A^{\mathrm{T}} + Q - \mathcal{K}C\mathbb{P}A^{\mathrm{T}} - A\mathbb{P}C^{\mathrm{T}}\mathcal{K}^{\mathrm{T}} && (2.78) \\
0 &= A\mathbb{P}C^{\mathrm{T}} && (2.79)
\end{aligned}$$

Assume that A is invertible and C has LI columns. Then (2.79) implies $\mathbb{P} = 0$, and then (2.78) and (2.76) imply $Q = 0$ and $R = 0$, respectively.

Assume A is not invertible. Let $z \neq 0$ such that $Az = 0$. Set $\mathbb{P} = zz^{\mathrm{T}} \neq 0$, $Q = \mathbb{P}$, and $R = -C\mathbb{P}C^{\mathrm{T}}$. Then (2.75)–(2.79) are satisfied. Thus $\begin{bmatrix} Q_{\mathrm{ss}} \\ R_{\mathrm{ss}} \end{bmatrix}$ is a nonzero element of null $\tilde{\mathbb{U}}$.

Assume C has LD columns. Let $z \neq 0$ such that $Cz = 0$. Set $\mathbb{P} = zz^{\mathrm{T}} \neq 0$, $Q = \mathbb{P} - A\mathbb{P}A^{\mathrm{T}}$, and $R = 0$. Then (2.75)–(2.79) are satisfied. Moreover, $\mathbb{P} \neq 0 \implies Q \neq 0$ (recall (2.19)). Thus $\left[\begin{smallmatrix} Q_{\mathrm{ss}} \\ R_{\mathrm{ss}} \end{smallmatrix}\right]$ is a nonzero element of null $\tilde{\mathbb{U}}$. ■

Remark 2.11. If $S = 0$ and A is invertible, then (Φ, Θ) uniquely determines L. Observe that $S = 0$ implies $AL = K$, and then A invertible implies $L = A^{-1}K$, and (Φ, Θ) uniquely determines K (recall Theorem 2.8). Furthermore, if additionally $H = I_p$, then (Φ, Θ) also uniquely determines R. Observe that $S = 0$ and A invertible together imply $\mathbb{P}C^{\mathrm{T}} = A^{-1}(\Theta + \mathcal{K}\Phi)$. Then $H = I_p$ implies $R = \Phi - C\mathbb{P}C^{\mathrm{T}} = \Phi - CA^{-1}(\Theta + \mathcal{K}\Phi)$. To illustrate these points, consider the system from Remark 2.4 with $\mathcal{K} = 0_{n,p}$ and the set of of covariance matrices $(Q^{\#}, R^{\#}, S^{\#})$ from Remark 2.7. Then $Q^{\#} \neq Q^{\star}$, but $R^{\#} = R^{\star}$, $\tilde{f}_{\Phi}(Q^{\#}, R^{\#}) = \tilde{f}_{\Phi}(Q^{\star}, R^{\star})$, $\tilde{f}_{\Theta}(Q^{\#}, R^{\#}) = \tilde{f}_{\Theta}(Q^{\star}, R^{\star})$, $\tilde{f}_{K}(Q^{\#}, R^{\#}) = \tilde{f}_{K}(Q^{\star}, R^{\star})$, and $\tilde{f}_{L}(Q^{\#}, R^{\#}) = \tilde{f}_{L}(Q^{\star}, R^{\star})$.

Appendices

2.A Miscellaneous results

Lemma 2.1. *If a matrix A is stable, then $A \otimes A$ is also stable.*

Proof. This result follows from

$$\operatorname{eig}(A \otimes A) = \{\lambda\mu \ : \ \lambda \in \operatorname{eig} A, \ \mu \in \operatorname{eig} A\}$$

See Magnus and Neudecker (2019, Theorem 2.1) for proof of this fact. ■

Lemma 2.2. *For a matrix A,* $\operatorname{eig}(I - A) = \{1 - \lambda : \lambda \in \operatorname{eig} A\}$.

Proof. Let $\lambda \in \operatorname{eig} A$. Then $Ax = \lambda x$ for some x. Thus

$$(I - A)x = x - Ax = x - \lambda x = (1 - \lambda)x$$

so $1 - \lambda \in \operatorname{eig}(I - A)$. Now suppose $\lambda \notin \operatorname{eig} A$. Then $Ax \neq \lambda x$ for all $x \neq 0$. So for all $x \neq 0$, we have

$$(I - A)x = x - Ax \neq x - \lambda x = (1 - \lambda)x$$

so $1 - \lambda \notin \operatorname{eig}(I - A)$. ■

Corollary 2.7. *If a matrix A is stable, then $I - A$ is invertible.*

Proof. If A is stable, it cannot have 1 as an eigenvalue, and therefore $I - A$ cannot have 0 as an eigenvalue. Hence $I - A$ is invertible. ■

Corollary 2.8. *If a matrix A is stable, then $I - (A \otimes A)$ is invertible.*

Proof. This result follows from Lemma 2.1 and Corollary 2.7. ■

Lemma 2.3. *If a matrix A is stable, then*

$$(I - A)^{-1} = \lim_{k \to \infty} S_k \qquad S_k := \sum_{i=0}^{k} A^i$$

Proof. We have

$$S_k - AS_k = (I - A)S_k = \sum_{i=0}^{k} A^i - \sum_{i=1}^{k+1} A^i = (I - A^{k+1})$$

Because $I - A$ is invertible (Corollary 2.7), we get $S_k = (I - A)^{-1}(I - A^{k+1})$. The result follows by taking the limit of this equation as $k \to \infty$. ■

Theorem 2.13. *Let A be a stable matrix and b be a column vector with the same number of rows as A. Then*

1. *$x^* = (I - A)^{-1}b$ is the unique solution to the equation $x = Ax + b$.*
2. *Recursively define a sequence by $x_{k+1} = Ax_k + b$. For any initial condition x_0, $x_k \to x^*$.*

Proof.

1. Since $I - A$ is invertible (Corollary 2.7), $(I - A)x = b \iff x = x^*$.
2. By induction, we have that

$$x_k = A^k x_0 + \Big(\sum_{i=0}^{k-1} A^i\Big)b$$

The result then follows by taking the limit of this equation and applying Lemma 2.3. ■

Corollary 2.9. *Let A be a stable matrix and Q be a square matrix of the same size as A. Then*

1. *The discrete Lyapunov equation $P = APA^{\mathrm{T}} + Q$ has a unique solution P^* where $P_{\mathrm{s}}^* = (I - A \otimes A)^{-1} Q_{\mathrm{s}}$.*[4]
2. *Recursively define a sequence by $P_{k+1} = AP_kA^{\mathrm{T}} + Q$. For any initial condition P_0, $P_k \to P^*$. Moreover, if Q is symmetric, then P^* is symmetric. Similarly, $Q \geq 0 \implies P^* \geq 0$ and $Q > 0 \implies P^* > 0$.*

Proof. The results follow from vectorizing and rearranging the discrete Lyapunov equation and applying Lemma 2.1 and Theorem 2.13. The point that Q symmetric implies P^* symmetric follows by induction: if P_0 is symmetric, then Q symmetric implies that P_k is symmetric for all $k > 0$. Thus the limit P^* must be symmetric as well. The point regarding positive semidefiniteness follows by the same reasoning. The point regarding positive definiteness then follows by noting that $P^* \geq 0$ and $Q > 0 \implies P^* = AP^*A^{\mathrm{T}} + Q > 0$. ∎

Theorem 2.14. *Suppose*

$$\begin{bmatrix} A & B \\ B^{\mathrm{T}} & C \end{bmatrix} \geq 0$$

Then $B = AX = YC$ for some matrices X and Y.

Proof. We decompose the matrix using the Cholesky decomposition:

$$\begin{bmatrix} A & B \\ B^{\mathrm{T}} & C \end{bmatrix} = \begin{bmatrix} L \\ M \end{bmatrix} \begin{bmatrix} L^{\mathrm{T}} & M^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} LL^{\mathrm{T}} & LM^{\mathrm{T}} \\ ML^{\mathrm{T}} & MM^{\mathrm{T}} \end{bmatrix}$$

Because $\operatorname{col} LL^{\mathrm{T}} = \operatorname{col} L$ (see Lemma A.1 in Appendix A), $B = LM^{\mathrm{T}} = LL^{\mathrm{T}}X = AX$ for some matrix X. The general solution to this equation is $X = A^{+}B + (I - A^{+}A)Z$, where Z is an arbitrary matrix (Magnus and Neudecker, 2019, Theorem 2.13, page 43). Similarly, $B^{\mathrm{T}} = ML^{\mathrm{T}} = MM^{\mathrm{T}}Y^{\mathrm{T}} = CY^{\mathrm{T}}$ for some matrix Y. Transposing this equation gives $B = YC$, for which the general solution is $Y = BC^{+} + W(I - CC^{+})$, where W is an arbitrary matrix. ∎

Corollary 2.10. $AA^{+}B = BC^{+}C = B$.

Proof. $AA^{+}B = AA^{+}AX = AX = B$. Similarly, $BC^{+}C = YCC^{+}C = YC = B$. ∎

[4] Moreover, if Q is positive semidefinite and (A, Q) is controllable, then P^* is positive definite. This fact is unnecessary for the work in this dissertation and thus the proof is omitted. A proof may be found in Anderson and Moore (1979, Section 4.2).

2.B Alternate transformation when A is stable

For this section, consider the process system with inputs, (1.1) and (1.2), and assume that A is stable. In this case (x_k) and (y_k) are not necessarily wide-sense stationary as $k \to \infty$. That is, (2.20)–(2.24) no longer hold. However, we can calculate transformed outputs that are approximately generated by a system of the same form as (1.3) and (1.4). Define

$$\tilde{x}_k := \begin{cases} 0_{n\times 1}, & k = 0 \\ \sum_{j=0}^{k-1} A^{k-j-1} G w_j, & k > 0 \end{cases} \qquad \tilde{u}_k := \begin{cases} 0_{n\times 1}, & k = 0 \\ \sum_{j=0}^{k-1} A^{k-j-1} B u_j, & k > 0 \end{cases}$$

Observe that

$$\tilde{x}_{k+1} = A\tilde{x}_k + Gw_k \tag{2.80}$$

A non-recursive equation for x_k is

$$x_k = A^k x_0 + \tilde{u}_k + \tilde{x}_k, \quad k \geq 0$$

Because A is stable, $A^k \to 0$ as $k \to \infty$. Thus we make the approximation $A^k = 0$ for $k \geq M$ for some integer M. Then

$$x_k = \tilde{u}_k + \tilde{x}_k, \quad k \geq M$$

Now define the transformed output:

$$\begin{aligned} \tilde{y}_k &:= y_k - C\tilde{u}_k - Du_k \\ &= C\tilde{x}_k + y_k - Cx_k - Du_k \\ &= C\tilde{x}_k + Hv_k \end{aligned} \tag{2.81}$$

Together (2.80) and (2.81) form a system in the process form without inputs. The state transition matrix is the same A, which we assumed was stable. Therefore (2.20)–(2.24) hold for $(\tilde{x}_k)$ and $(\tilde{y}_k)$ as $k \to \infty$.

We adore chaos because we love to produce order.

M. C. Escher

CHAPTER 3

ALS PROBLEM FORMULATIONS AND UNIQUENESS CONDITIONS

Continuing along the lines of Chapter 2, this chapter presents the autocovariance least squares (ALS) technique for covariance matrix estimation. The derivation of the ALS problem parallels that which has been given in previous publications, such as Odelson et al. (2006b) and Rajamani and Rawlings (2009). Most authors assume $S = 0$ and pose ALS problems to estimate (Q, R). This form of the ALS problem is considered along with three others: 1) (Q, R, S),[1] 2) (Φ, Θ), and 3) $(\mathbb{\Phi}, \mathbb{\Theta})$. The (Q, R, S) and $(Q, R, 0)$ problems may be posed for either the process system without inputs and A stable or for the $\mathcal{K}$-error system. The (Φ, Θ) problem may be posed only for the process system without inputs and A stable. The $(\mathbb{\Phi}, \mathbb{\Theta})$ problem may be posed only for the $\mathcal{K}$-error system.

This chapter is organized as follows:

- Section 3.1 derives the unconstrained ALS problems.
- Section 3.2 gives the uniqueness conditions for each form of the ALS problem under the assumptions $G = I_n$ and $H = I_p$. It is demonstrated that dropping these assumptions can alter the identifiability of $(Q, R, 0)$ and (Q, R, S), and

[1] To the author's knowledge, Åkesson et al. (2008) are the only authors to have previously developed an ALS method for estimating (Q, R, S), but they do not discuss uniqueness conditions for the problem, as is done here in Section 3.2.

that, in general, $(\Phi, \Theta)/(\mathbb{\Phi}, \mathbb{\Theta})$ have the least stringent and (Q, R, S) the most stringent requirements to be identified uniquely, with $(Q, R, 0)$ in the middle.

- Section 3.3 discusses the relationship between ALS estimates $(\hat{Q}, \hat{R}, \hat{S})$ and $(\hat{\Phi}, \hat{\Theta})$ for the process system without inputs and A stable, and between estimates $(\hat{Q}, \hat{R}, \hat{S})$ and $(\hat{\mathbb{\Phi}}, \hat{\mathbb{\Theta}})$ for the $\mathcal{K}$-error system.
- Section 3.4 presents the constrained versions of the ALS problems.

The chapter also contains five appendices: 3.A, 3.B, 3.C, 3.D, and 3.E.

3.1 Derivation of ALS problems

We begin with another assumption:

Assumption 7. The outputs $y_{0:T_f}$ and the corresponding $\mathcal{K}$-innovations $\mathbb{z}_{0:T_f}$ have the steady-state autocovariances given in (2.21)–(2.24) and (2.28)–(2.31), respectively.

In practice, one may need to exclude an initial window of $y_{0:T_f}$ or $\mathbb{z}_{0:T_f}$ from analysis to ensure that Assumption 7 holds. We examine the validity of Assumption 7 in Appendix 3.B.

The autocovariance at lag j is defined as

$$\Lambda_j \coloneqq \mathrm{E}[y_{k+j} y_k^{\mathrm{T}}] \qquad \mathbb{\Lambda}_j \coloneqq \mathrm{E}[\mathbb{z}_{k+j} \mathbb{z}_k^{\mathrm{T}}]$$

Note that $\Lambda_{-j} = \Lambda_j^{\mathrm{T}}$ and $\mathbb{\Lambda}_{-j} = \mathbb{\Lambda}_j^{\mathrm{T}}$. The autocovariance matrix with N lags is defined as follows, and it can be expressed as a linear function of (P, R, S) or of (Φ, Θ):

$$\begin{aligned}
\Psi \coloneqq \begin{bmatrix} \Lambda_0 \\ \Lambda_1 \\ \Lambda_2 \\ \vdots \\ \Lambda_{N-1} \end{bmatrix} &= \begin{bmatrix} CPC^{\mathrm{T}} + HRH^{\mathrm{T}} \\ C(APC^{\mathrm{T}} + GSH^{\mathrm{T}}) \\ CA(APC^{\mathrm{T}} + GSH^{\mathrm{T}}) \\ \vdots \\ CA^{N-2}(APC^{\mathrm{T}} + GSH^{\mathrm{T}}) \end{bmatrix} = \begin{bmatrix} \Phi \\ C\Theta \\ CA\Theta \\ \vdots \\ CA^{N-2}\Theta \end{bmatrix} \\
&= \mathcal{O}_{1,N} P C^{\mathrm{T}} + \Gamma H R H^{\mathrm{T}} + \mathcal{O}_{2,N} G S H^{\mathrm{T}} = \Gamma \Phi + \mathcal{O}_{2,N} \Theta \qquad (3.1)
\end{aligned}$$

$$\mathcal{O}_{1,N} \coloneqq \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{N-1} \end{bmatrix} \qquad \mathcal{O}_{2,N} \coloneqq \begin{bmatrix} 0_{p,n} \\ C \\ CA \\ \vdots \\ CA^{N-2} \end{bmatrix} \qquad \Gamma \coloneqq \begin{bmatrix} I_p \\ 0_{p,p} \\ 0_{p,p} \\ \vdots \\ 0_{p,p} \end{bmatrix}$$

Note the identity $\mathcal{O}_{1,N} = \mathcal{O}_{2,N}A + \Gamma C$. For the $\mathcal{K}$-error system, the analogous equations are

$$\Psi := \begin{bmatrix} \mathbb{\Lambda}_0 \\ \mathbb{\Lambda}_1 \\ \mathbb{\Lambda}_2 \\ \vdots \\ \mathbb{\Lambda}_{N-1} \end{bmatrix} = \begin{bmatrix} C\mathbb{P}C^{\mathrm{T}} + HRH^{\mathrm{T}} \\ C(\mathbb{A}\mathbb{P}C^{\mathrm{T}} - \mathcal{K}HRH^{\mathrm{T}} + GSH^{\mathrm{T}}) \\ C\mathbb{A}(\mathbb{A}\mathbb{P}C^{\mathrm{T}} - \mathcal{K}HRH^{\mathrm{T}} + GSH^{\mathrm{T}}) \\ \vdots \\ C\mathbb{A}^{N-2}(\mathbb{A}\mathbb{P}C^{\mathrm{T}} - \mathcal{K}HRH^{\mathrm{T}} + GSH^{\mathrm{T}}) \end{bmatrix} = \begin{bmatrix} \mathbb{\Phi} \\ C\mathbb{\Theta} \\ C\mathbb{A}\mathbb{\Theta} \\ \vdots \\ C\mathbb{A}^{N-2}\mathbb{\Theta} \end{bmatrix}$$
$$= \mathbb{O}_{1,N}\mathbb{P}C^{\mathrm{T}} + \mathbb{\Gamma}HRH^{\mathrm{T}} + \mathbb{O}_{2,N}GSH^{\mathrm{T}} = \Gamma\mathbb{\Phi} + \mathbb{O}_{2,N}\mathbb{\Theta}$$
$$\mathbb{O}_{1,N} := \begin{bmatrix} C \\ C\mathbb{A} \\ C\mathbb{A}^2 \\ \vdots \\ C\mathbb{A}^{N-1} \end{bmatrix} \qquad \mathbb{O}_{2,N} := \begin{bmatrix} 0_{p,n} \\ C \\ C\mathbb{A} \\ \vdots \\ C\mathbb{A}^{N-2} \end{bmatrix} \qquad \mathbb{\Gamma} := \Gamma - \mathbb{O}_{2,N}\mathcal{K}$$

Note the identity $\mathbb{O}_{1,N} = \mathbb{O}_{2,N}\mathbb{A} + \Gamma C$. Also note that N must be chosen such that $N \leq T_f + 1$. The vectorized form of Ψ in terms of Q, R, and S is

$$\begin{aligned} \Psi_{\mathrm{s}} &= (C \otimes \mathcal{O}_{1,N})P_{\mathrm{s}} + (H \otimes \Gamma H)R_{\mathrm{s}} + (H \otimes \mathcal{O}_{2,N}G)S_{\mathrm{s}} \\ &= (C \otimes \mathcal{O}_{1,N})F(G \otimes G)Q_{\mathrm{s}} + (H \otimes \Gamma H)R_{\mathrm{s}} + (H \otimes \mathcal{O}_{2,N}G)S_{\mathrm{s}} \\ &= (C \otimes \mathcal{O}_{1,N})F(G \otimes G)\mathscr{D}_g Q_{\mathrm{ss}} + (H \otimes \Gamma H)\mathscr{D}_h R_{\mathrm{ss}} + (H \otimes \mathcal{O}_{2,N}G)S_{\mathrm{s}} \\ &= X_Q Q_{\mathrm{ss}} + X_R R_{\mathrm{ss}} + X_S S_{\mathrm{s}} \end{aligned}$$
$$X_Q := (C \otimes \mathcal{O}_{1,N})F(G \otimes G)\mathscr{D}_g$$
$$X_R := (H \otimes \Gamma H)\mathscr{D}_h \qquad X_S := (H \otimes \mathcal{O}_{2,N}G)$$

The vectorized form of Ψ in terms of Φ and Θ is

$$\begin{aligned} \Psi_{\mathrm{s}} &= (I_p \otimes \Gamma)\Phi_{\mathrm{s}} + (I_p \otimes \mathcal{O}_{2,N})\Theta_{\mathrm{s}} \\ &= (I_p \otimes \Gamma)\mathscr{D}_p\Phi_{\mathrm{ss}} + (I_p \otimes \mathcal{O}_{2,N})\Theta_{\mathrm{s}} \\ &= X_\Phi\Phi_{\mathrm{ss}} + X_\Theta\Theta_{\mathrm{s}} \end{aligned}$$
$$X_\Phi := (I_p \otimes \Gamma)\mathscr{D}_p \qquad X_\Theta := (I_p \otimes \mathcal{O}_{2,N})$$

The vectorized form of Ψ in terms of Q, R, and S is

$$\begin{aligned} \Psi_{\mathrm{s}} &= (C \otimes \mathbb{O}_{1,N})\mathbb{P}_{\mathrm{s}} + (H \otimes \mathbb{\Gamma}H)R_{\mathrm{s}} + (H \otimes \mathbb{O}_{2,N}G)S_{\mathrm{s}} \\ &= (C \otimes \mathbb{O}_{1,N})\mathbb{F}[(G \otimes G)Q_{\mathrm{s}} + (\mathcal{K}H \otimes \mathcal{K}H)R_{\mathrm{s}} - (I_{n^2} + \mathscr{K}_n)(\mathcal{K}H \otimes G)S_{\mathrm{s}}] \\ &\quad + (H \otimes \mathbb{\Gamma}H)R_{\mathrm{s}} + (H \otimes \mathbb{O}_{2,N}G)S_{\mathrm{s}} \end{aligned}$$

$$
\begin{aligned}
&= (C \otimes \mathbb{O}_{1,N})\mathbb{F}(G \otimes G)Q_{\mathrm{s}} \\
&\quad + [(C \otimes \mathbb{O}_{1,N})\mathbb{F}(\mathcal{K}H \otimes \mathcal{K}H) + (H \otimes \mathbb{\Gamma} H)]R_{\mathrm{s}} \\
&\quad + [-(C \otimes \mathbb{O}_{1,N})\mathbb{F}(I_{n^2} + \mathscr{K}_n)(\mathcal{K}H \otimes G) + (H \otimes \mathbb{O}_{2,N}G)]S_{\mathrm{s}} \\
&= (C \otimes \mathbb{O}_{1,N})\mathbb{F}(G \otimes G)\mathscr{D}_g Q_{\mathrm{ss}} \\
&\quad + [(C \otimes \mathbb{O}_{1,N})\mathbb{F}(\mathcal{K}H \otimes \mathcal{K}H) + (H \otimes \mathbb{\Gamma} H)]\mathscr{D}_h R_{\mathrm{ss}} \\
&\quad + [-(C \otimes \mathbb{O}_{1,N})\mathbb{F}(I_{n^2} + \mathscr{K}_n)(\mathcal{K}H \otimes G) + (H \otimes \mathbb{O}_{2,N}G)]S_{\mathrm{s}} \\
&= \mathbb{X}_Q Q_{\mathrm{ss}} + \mathbb{X}_R R_{\mathrm{ss}} + \mathbb{X}_S S_{\mathrm{s}} \\
&\qquad\qquad \mathbb{X}_Q \coloneqq (C \otimes \mathbb{O}_{1,N})\mathbb{F}(G \otimes G)\mathscr{D}_g \\
&\qquad\qquad \mathbb{X}_R \coloneqq [(C \otimes \mathbb{O}_{1,N})\mathbb{F}(\mathcal{K}H \otimes \mathcal{K}H) + (H \otimes \mathbb{\Gamma} H)]\mathscr{D}_h \\
&\qquad\qquad \mathbb{X}_S \coloneqq -(C \otimes \mathbb{O}_{1,N})\mathbb{F}(I_{n^2} + \mathscr{K}_n)(\mathcal{K}H \otimes G) + (H \otimes \mathbb{O}_{2,N}G)
\end{aligned}
$$

The vectorized form of $\mathbb{\Psi}$ in terms of $\mathbb{\Phi}$ and $\mathbb{\Theta}$ is

$$
\begin{aligned}
\mathbb{\Psi}_{\mathrm{s}} &= (I_p \otimes \mathbb{\Gamma})\mathbb{\Phi}_{\mathrm{s}} + (I_p \otimes \mathbb{O}_{2,N})\mathbb{\Theta}_{\mathrm{s}} \\
&= (I_p \otimes \mathbb{\Gamma})\mathscr{D}_p \mathbb{\Phi}_{\mathrm{ss}} + (I_p \otimes \mathbb{O}_{2,N})\mathbb{\Theta}_{\mathrm{s}} \\
&= \mathbb{X}_{\mathbb{\Phi}} \mathbb{\Phi}_{\mathrm{ss}} + \mathbb{X}_{\mathbb{\Theta}} \mathbb{\Theta}_{\mathrm{s}} \\
\mathbb{X}_{\mathbb{\Phi}} &\coloneqq (I_p \otimes \mathbb{\Gamma})\mathscr{D}_p \qquad \mathbb{X}_{\mathbb{\Theta}} \coloneqq (I_p \otimes \mathbb{O}_{2,N})
\end{aligned}
$$

Removing redundant elements gives

$$
b \coloneqq \Psi_{\mathrm{ss}} = \mathscr{D}^{+}_{Np,p}\Psi_{\mathrm{s}} = X_{QRS}\begin{bmatrix} Q_{\mathrm{ss}} \\ R_{\mathrm{ss}} \\ S_{\mathrm{s}} \end{bmatrix} = X_{\Phi\Theta}\begin{bmatrix} \Phi_{\mathrm{ss}} \\ \Theta_{\mathrm{s}} \end{bmatrix} \tag{3.2}
$$

$$
\mathbb{b} \coloneqq \mathbb{\Psi}_{\mathrm{ss}} = \mathscr{D}^{+}_{Np,p}\mathbb{\Psi}_{\mathrm{s}} = \mathbb{X}_{QRS}\begin{bmatrix} Q_{\mathrm{ss}} \\ R_{\mathrm{ss}} \\ S_{\mathrm{s}} \end{bmatrix} = \mathbb{X}_{\mathbb{\Phi}\mathbb{\Theta}}\begin{bmatrix} \mathbb{\Phi}_{\mathrm{ss}} \\ \mathbb{\Theta}_{\mathrm{s}} \end{bmatrix} \tag{3.3}
$$

$$
\begin{aligned}
X_{QRS} &\coloneqq \mathscr{D}^{+}_{Np,p}\begin{bmatrix} X_Q & X_R & X_S \end{bmatrix} & \mathbb{X}_{QRS} &\coloneqq \mathscr{D}^{+}_{Np,p}\begin{bmatrix} \mathbb{X}_Q & \mathbb{X}_R & \mathbb{X}_S \end{bmatrix} \\
X_{QR} &\coloneqq \mathscr{D}^{+}_{Np,p}\begin{bmatrix} X_Q & X_R \end{bmatrix} & \mathbb{X}_{QR} &\coloneqq \mathscr{D}^{+}_{Np,p}\begin{bmatrix} \mathbb{X}_Q & \mathbb{X}_R \end{bmatrix} \\
X_{\Phi\Theta} &\coloneqq \mathscr{D}^{+}_{Np,p}\begin{bmatrix} X_\Phi & X_\Theta \end{bmatrix} & \mathbb{X}_{\mathbb{\Phi}\mathbb{\Theta}} &\coloneqq \mathscr{D}^{+}_{Np,p}\begin{bmatrix} \mathbb{X}_{\mathbb{\Phi}} & \mathbb{X}_{\mathbb{\Theta}} \end{bmatrix}
\end{aligned}
$$

Theorem 3.1. *The following equations hold:*

$$
X_{QRS} = X_{\Phi\Theta}U \qquad \mathbb{X}_{QRS} = \mathbb{X}_{\mathbb{\Phi}\mathbb{\Theta}}\mathbb{U}
$$

Proof of Theorem 3.1 is provided in Appendix 3.C.

The vectors b and $\mathbb{b}$ must be estimated from the output data. In light of Assumption 7, we have that

$$\hat{\Lambda}_j \coloneqq \frac{1}{T_f - j + 1} \sum_{i=0}^{T_f - j} y_{i+j} y_i^{\mathrm{T}} \qquad \hat{\mathbb{\Lambda}}_j \coloneqq \frac{1}{T_f - j + 1} \sum_{i=0}^{T_f - j} \mathbb{z}_{i+j} \mathbb{z}_i^{\mathrm{T}} \tag{3.4}$$

are unbiased estimators of Λ_j and $\mathbb{\Lambda}_j$, respectively, so

$$\hat{\Psi} \coloneqq \begin{bmatrix} \hat{\Lambda}_0 \\ \hat{\Lambda}_1 \\ \vdots \\ \hat{\Lambda}_{N-1} \end{bmatrix} \qquad \hat{\mathbb{\Psi}} \coloneqq \begin{bmatrix} \hat{\Lambda}_0 \\ \hat{\Lambda}_1 \\ \vdots \\ \hat{\Lambda}_{N-1} \end{bmatrix}$$

are unbiased estimators of Ψ and $\mathbb{\Psi}$, respectively. Therefore

$$\hat{b} \coloneqq \hat{\Psi}_{\mathrm{ss}} \qquad \hat{\mathbb{b}} \coloneqq \hat{\mathbb{\Psi}}_{\mathrm{ss}} \tag{3.5}$$

are unbiased estimators of b and $\mathbb{b}$, respectively.

Equation (3.2) can be rewritten as a linear model in terms of $\hat{b}$ and $e \coloneqq \hat{b} - b$, and (3.3) can be rewritten as a linear model in terms of $\hat{\mathbb{b}}$ and $\mathbb{e} = \hat{\mathbb{b}} - \mathbb{b}$:

$$\begin{aligned} \hat{b} &= X_{QRS} \begin{bmatrix} Q_{\mathrm{ss}} \\ R_{\mathrm{ss}} \\ S_{\mathrm{s}} \end{bmatrix} + e = X_{\Phi\Theta} \begin{bmatrix} \Phi_{\mathrm{ss}} \\ \Theta_{\mathrm{s}} \end{bmatrix} + e \\ \hat{\mathbb{b}} &= \mathbb{X}_{QRS} \begin{bmatrix} Q_{\mathrm{ss}} \\ R_{\mathrm{ss}} \\ S_{\mathrm{s}} \end{bmatrix} + \mathbb{e} = \mathbb{X}_{\Phi\Theta} \begin{bmatrix} \Phi_{\mathrm{ss}} \\ \Theta_{\mathrm{s}} \end{bmatrix} + \mathbb{e} \end{aligned} \tag{3.6}$$

Therefore, using results from Appendix A, the minimum variance affine unbiased estimate of (Q, R, S), $(Q, R, 0)$, (Φ, Θ), or $(\mathbb{\Phi}, \mathbb{\Theta})$ from $\hat{b}$ or $\hat{\mathbb{b}}$ is the solution to a weighted linear least squares problem. Specifically, applying Corollary A.2 gives the six different forms of the ALS problem

$$\min_{Q,R,S} \tfrac{1}{2} \left\| X_{QRS} \begin{bmatrix} Q_{\mathrm{ss}} \\ R_{\mathrm{ss}} \\ S_{\mathrm{s}} \end{bmatrix} - \hat{b} \right\|_{W_{QRS}}^2 \tag{3.7}$$

$$\min_{Q,R} \tfrac{1}{2} \left\| X_{QR} \begin{bmatrix} Q_{\mathrm{ss}} \\ R_{\mathrm{ss}} \end{bmatrix} - \hat{b} \right\|_{W_{QR}}^2 \tag{3.8}$$

$$\min_{\Phi,\Theta} \tfrac{1}{2} \left\| X_{\Phi\Theta} \begin{bmatrix} \Phi_{\mathrm{ss}} \\ \Theta_{\mathrm{s}} \end{bmatrix} - \hat{b} \right\|_{W_{\Phi\Theta}}^2 \tag{3.9}$$

$$\min_{Q,R,S} \tfrac{1}{2} \left\| \mathbb{X}_{QRS} \begin{bmatrix} Q_{\mathrm{ss}} \\ R_{\mathrm{ss}} \\ S_{\mathrm{s}} \end{bmatrix} - \hat{\mathbb{b}} \right\|_{\mathbb{W}_{QRS}}^2 \tag{3.10}$$

$$\min_{Q,R} \tfrac{1}{2} \left\| \mathbb{X}_{QR} \begin{bmatrix} Q_{\mathrm{ss}} \\ R_{\mathrm{ss}} \end{bmatrix} - \hat{\mathbb{b}} \right\|_{\mathbb{W}_{QR}}^2 \tag{3.11}$$

$$\min_{\mathbb{\Phi},\mathbb{\Theta}} \tfrac{1}{2} \left\| \mathbb{X}_{\mathbb{\Phi}\mathbb{\Theta}} \begin{bmatrix} \mathbb{\Phi}_{\mathrm{ss}} \\ \mathbb{\Theta}_{\mathrm{s}} \end{bmatrix} - \hat{\mathbb{b}} \right\|_{\mathbb{W}_{\mathbb{\Phi}\mathbb{\Theta}}}^2 \tag{3.12}$$

where $W_\Omega := (P_{\hat{b}} + X_\Omega X_\Omega^{\mathrm{T}})^+$ and $\mathbb{W}_\Omega := (P_{\hat{\mathbb{b}}} + \mathbb{X}_\Omega \mathbb{X}_\Omega^{\mathrm{T}})^+$, and $P_{\hat{b}} := \mathrm{cov}(\hat{b})$ and $P_{\hat{\mathbb{b}}} := \mathrm{cov}(\hat{\mathbb{b}})$, and Ω refers to any of (Q, R, S), $(Q, R, 0)$, (Φ, Θ), or $(\Phi, \mathbb{\Theta})$.

The optimal weighting matrix for the ALS problem is generally not available, as $P_{\hat{b}}$ and $P_{\hat{\mathbb{b}}}$ depend on the true values of the sought parameters. Zagrobelny and Rawlings (2015b) were the first to implement a method for estimating the ALS weighting matrix from data. The latest results on ALS weight estimation are available in Chapters 4, 6 and 7 of this dissertation and in Arnold and Rawlings (2020b). For commentary on the practical aspects of solving ALS problems and numerical examples, we refer the reader to Chapter 7 of this dissertation and to Odelson et al. (2006b), Rajamani and Rawlings (2009), Zagrobelny and Rawlings (2015b), and Arnold and Rawlings (2020b).

3.2 Uniqueness conditions for ALS problems

We next give necessary and sufficient conditions for the constant matrices X_{QRS}, $\mathbb{X}_{QRS}$, X_{QR}, $\mathbb{X}_{QR}$, $X_{\Phi\Theta}$, and $\mathbb{X}_{\Phi\mathbb{\Theta}}$ to have LI columns, which from Corollary A.2 is necessary and sufficient for the respective least squares problems (3.7)–(3.12) to have unique solutions.

Theorem 3.2. *For the following, when considering the process system, assume A is stable. When considering the $\mathcal{K}$-error system, assume $\mathbb{A} = A - \mathcal{K}C$ is stable.*

1. *If $G = I_n$ and $H = I_p$, then X_{QRS} and $\mathbb{X}_{QRS}$ have LD columns.*
2. *If $N \geq 2$, $G = I_n$, and $H = I_p$, then X_{QR} and $\mathbb{X}_{QR}$ have LI columns if and only if C has LI columns and A is invertible.*[2]
3. *If $N \geq n_{\mathrm{obs}} + 1$, where n_{obs} is the observability index*[3] *of (A, C), then $X_{\Phi\Theta}$ and $\mathbb{X}_{\Phi\mathbb{\Theta}}$ have LI columns if and only if (A, C) is observable.*

Proof of Theorem 3.2 is provided in Appendix 3.D. Theorem 3.2 is consistent with Corollaries 2.3–2.6 in the sense that if $G = I_n$ and $H = I_p$, then (Q, R, S)

[2] Rajamani and Rawlings (2009) mention that if A is singular, a similarity transformation can be used to transform the model and eliminate the zero-eigenvalue states. This assertion is true if and only if A is nondefective.

[3] We define the observability index of (A, C) as the smallest natural number n_{obs} such that $\mathrm{rank}\, \mathcal{O}_{1,n_{\mathrm{obs}}} = \mathrm{rank}\, \mathcal{O}_{1,n}$.

and $(Q, R, 0)$ have unique estimates under the same conditions in which they are uniquely determined by (Φ, Θ) and $(\Phi, \mathbb{\Theta})$.

Rajamani and Rawlings (2009) and Zagrobelny and Rawlings (2015b) have previously studied ALS uniqueness conditions, but only for $(Q, R, 0)$ estimation. In particular, Zagrobelny and Rawlings (2015b, Theorem 1) prove something similar to the second part of Theorem 3.2. There are two differences in the version presented here. First, we have shown that A invertible is a necessary condition in addition to being one of the sufficient conditions. Second, (A, C) observable is not included as a necessary and sufficient condition, as this condition is redundant given that C having full column rank is a necessary condition and one of the sufficient conditions.

We have not established concise uniqueness conditions for the $(Q, R, 0)$ and (Q, R, S) ALS problems, (3.7), (3.8), (3.10), and (3.11), when $G \neq I_n$ and/or $H \neq I_p$. When $G \neq I_n$ and/or $H \neq I_p$, the easiest way to check for uniqueness is to numerically check if the columns of X_{QRS}, $\mathbb{X}_{QRS}$, X_{QR}, and $\mathbb{X}_{QR}$ are LI. It is known that the results of Theorem 3.2 do not apply if $G \neq I_n$ and/or $H \neq I_p$. For example, consider the following system:

$$A = \begin{bmatrix} 0.1 & 0 & 0 \\ 0 & 0.2 & 0 \\ 0 & 0 & 0.3 \end{bmatrix} \qquad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \qquad G = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad H = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

It is straightforward to verify that for this system, for $N \geq 2$, X_{QRS}, X_{QR}, $\mathbb{X}_{QRS}$ and $\mathbb{X}_{QR}$ (take $\mathcal{K} = 0_{n,p}$ for the latter two) have LI columns, which is not the case for the same system with $G = I_3$ instead.

The requirements to uniquely estimate (Φ, Θ) and $(\Phi, \mathbb{\Theta})$ from the sample autocovariance are significantly weaker than those for (Q, R, S) and $(Q, R, 0)$. With $G = I_n$ and $H = I_p$, there is no unique estimate for (Q, R, S), and there is a unique estimate of $(Q, R, 0)$ if and only if every state in the model is measured, which is often not the case in practice. On the other hand, there is a unique estimate of (Φ, Θ) and $(\Phi, \mathbb{\Theta})$ if and only if the system is observable, and one can always obtain an observable system model by passing to a minimal realization to remove any unobservable states.

3.3 Relationship between ALS estimates $(\hat{Q}, \hat{R}, \hat{S})$, $(\hat{\Phi}, \hat{\Theta})$, and $(\hat{\Phi}, \hat{\mathbb{\Theta}})$

The following theorems show that under mild assumptions, even if the ALS estimate $(\hat{Q}, \hat{R}, \hat{S})$ is not unique, any such estimate corresponds to the same $(\hat{\Phi}, \hat{\Theta})$ or $(\hat{\Phi}, \hat{\mathbb{\Theta}})$.

Theorem 3.3. *Assume A is stable, $X_{\Phi\Theta}$ has LI columns, U in (2.62) has LI rows, and that the weighting matrices for (3.7) and (3.9) are chosen as $W_{QRS} = (V + X_{QRS}X_{QRS}^{\mathrm{T}})^{+}$ and $W_{\Phi\Theta} = (V + X_{\Phi\Theta}X_{\Phi\Theta}^{\mathrm{T}})^{+}$, where V is an arbitrary positive semidefinite matrix. Then $f_{\Phi}(\hat{Q}, \hat{R}, \hat{S}) = \hat{\Phi}$ and $f_{\Theta}(\hat{Q}, \hat{R}, \hat{S}) = \hat{\Theta}$, where $(\hat{Q}, \hat{R}, \hat{S})$ is any solution to (3.7) and $(\hat{\Phi}, \hat{\Theta})$ is the unique solution to (3.9).*

Theorem 3.4. *Assume $\mathbb{A}$ is stable, $\mathbb{X}_{\Phi\mathbb{\Theta}}$ has LI columns, $\mathbb{U}$ in (2.70) has LI rows, and that the weighting matrices for (3.10) and (3.12) are chosen as $\mathbb{W}_{QRS} = (\mathbb{V} + \mathbb{X}_{QRS}\mathbb{X}_{QRS}^{\mathrm{T}})^{+}$ and $\mathbb{W}_{\Phi\mathbb{\Theta}} = (\mathbb{V} + \mathbb{X}_{\Phi\mathbb{\Theta}}\mathbb{X}_{\Phi\mathbb{\Theta}}^{\mathrm{T}})^{+}$, where $\mathbb{V}$ is an arbitrary positive semidefinite matrix. Then $f_{\Phi}(\hat{Q}, \hat{R}, \hat{S}) = \hat{\Phi}$ and $f_{\mathbb{\Theta}}(\hat{Q}, \hat{R}, \hat{S}) = \hat{\mathbb{\Theta}}$, where $(\hat{Q}, \hat{R}, \hat{S})$ is any solution to (3.10) and $(\hat{\Phi}, \hat{\mathbb{\Theta}})$ is the unique solution to (3.12).*

Proof of Theorem 3.3 is provided in Appendix 3.E. Proof of Theorem 3.4 is not explicitly provided, but it is entirely analogous to that for Theorem 3.3.

Recall from Section 2.3 that (Φ, Θ) and $(\Phi, \mathbb{\Theta})$ uniquely determine K. Therefore Theorems 3.3 and 3.4 indicate that if the objective is to design a Kalman predictor (i.e., estimate K from the data), then it usually does not matter which form of the ALS problem is solved. That is, solving any of (3.7), (3.9), (3.10), and (3.12) will lead to the same estimated value of K. It is important to note that Theorems 3.3 and 3.4 hold for arbitrary positive semidefinite V and $\mathbb{V}$, respectively, since, as previously mentioned, the true values of $V = \mathrm{cov}(\hat{b})$ and $\mathbb{V} = \mathrm{cov}(\hat{\mathbb{b}})$ are generally not known a priori, as they depend on the true values of the sought parameters.

3.4 Constrained ALS problems

The ALS problems given in Section 3.1 do not include constraints. In practice, constraints may be included to ensure physically meaningful parameter estimates are obtained. The matrices Q and R are valid covariances if and only if

$$Q \geq 0 \quad \text{and} \quad R \geq 0 \tag{3.13}$$

The matrices Q, R, and S form a valid covariance if and only if

$$\begin{bmatrix} Q & S \\ S^{\mathrm{T}} & R \end{bmatrix} \geq 0 \tag{3.14}$$

Inequalities (3.13) and (3.14) are linear matrix inequalities. When such constraints are added to a linear least squares problem, it can be formulated as a standard form semidefinite programming (SDP) problem, for which reliable solvers exist. Interior point methods are commonly used to solve SDPs; see Boyd and Vandenberghe (2004) and Vandenberghe and Boyd (1996) as references on SDPs.

For the (Φ, Θ) and $(\mathbb{\Phi}, \mathbb{\Theta})$ ALS problems, (3.9) and (3.12) respectively, it is less obvious what constraints should be imposed to ensure physically meaningful estimates. The matrices Q, R, and S are the fundamental stochastic parameters driving the system, not Φ and Θ or $\mathbb{\Phi}$ and $\mathbb{\Theta}$. Therefore (Φ, Θ) is physically meaningful if and only if there exists (Q, R, S) satisfying inequality (3.14), $f_\Phi(Q, R, S) = \Phi$, and $f_\Theta(Q, R, S) = \Theta$. Similarly, $(\mathbb{\Phi}, \mathbb{\Theta})$ is physically meaningful if and only if there exists (Q, R, S) satisfying inequality (3.14), $f_{\mathbb{\Phi}}(Q, R, S) = \mathbb{\Phi}$, and $f_{\mathbb{\Theta}}(Q, R, S) = \mathbb{\Theta}$. Thus, we propose the following constrained least squares problem for (Φ, Θ):

$$\begin{aligned} \min_{\Phi,\Theta,Q,R,S,P} \ & \tfrac{1}{2} \left\| X_{\Phi\Theta} \begin{bmatrix} \Phi_{\mathrm{ss}} \\ \Theta_{\mathrm{s}} \end{bmatrix} - \hat{b} \right\|^2_{W_{\Phi\Theta}} \\ \text{s.t.} \ \ & \Phi = CPC^{\mathrm{T}} + HRH^{\mathrm{T}} \qquad \Theta = APC^{\mathrm{T}} + GSH^{\mathrm{T}} \\ & P = APA^{\mathrm{T}} + GQG^{\mathrm{T}} \qquad \begin{bmatrix} Q & S \\ S^{\mathrm{T}} & R \end{bmatrix} \geq 0 \end{aligned} \tag{3.15}$$

For $(\mathbb{\Phi}, \mathbb{\Theta})$, the corresponding problem is

$$\begin{aligned} \min_{\mathbb{\Phi},\mathbb{\Theta},Q,R,S,\mathbb{P}} \ & \tfrac{1}{2} \left\| \mathbb{X}_{\mathbb{\Phi}\mathbb{\Theta}} \begin{bmatrix} \mathbb{\Phi}_{\mathrm{ss}} \\ \mathbb{\Theta}_{\mathrm{s}} \end{bmatrix} - \hat{\mathbb{b}} \right\|^2_{\mathbb{W}_{\mathbb{\Phi}\mathbb{\Theta}}} \\ \text{s.t.} \ \ & \mathbb{\Phi} = C\mathbb{P}C^{\mathrm{T}} + HRH^{\mathrm{T}} \qquad \mathbb{\Theta} = \mathbb{A}\mathbb{P}C^{\mathrm{T}} + GSH^{\mathrm{T}} - \mathcal{K}HRH^{\mathrm{T}} \qquad \begin{bmatrix} Q & S \\ S^{\mathrm{T}} & R \end{bmatrix} \geq 0 \\ & \mathbb{P} = \mathbb{A}\mathbb{P}\mathbb{A}^{\mathrm{T}} + GQG^{\mathrm{T}} + \mathcal{K}HRH^{\mathrm{T}}\mathcal{K}^{\mathrm{T}} - GSH^{\mathrm{T}}\mathcal{K}^{\mathrm{T}} - \mathcal{K}HS^{\mathrm{T}}G^{\mathrm{T}} \end{aligned} \tag{3.16}$$

The constraints in problems (3.15) and (3.16) are linear matrix equalities and inequalities, so each problem can be formulated as a standard form SDP.

If $N \geq n_{\mathrm{obs}} + 1$ and (A, C) is observable, then $X_{\Phi\Theta}$ and $\mathbb{X}_{\mathbb{\Phi}\mathbb{\Theta}}$ have LI columns. Therefore if $W_{\Phi\Theta}$ and $\mathbb{W}_{\mathbb{\Phi}\mathbb{\Theta}}$ are chosen such that the Hessians $X^{\mathrm{T}}_{\Phi\Theta} W_{\Phi\Theta} X_{\Phi\Theta}$ and $\mathbb{X}^{\mathrm{T}}_{\mathbb{\Phi}\mathbb{\Theta}} \mathbb{W}_{\mathbb{\Phi}\mathbb{\Theta}} \mathbb{X}_{\mathbb{\Phi}\mathbb{\Theta}}$ are strictly positive definite, then the objective functions in (3.15)

and (3.16) are strongly convex if they are considered to be functions of only (Φ, Θ) and $(\mathbb{\Phi}, \mathbb{\Theta})$, respectively. However, the objective functions do not directly depend on decision variables Q, R, S, and $P/\mathbb{P}$, and therefore the optimization problems may not have unique solutions. In practice it may be desirable to add a small, strongly convex penalty to Q, R, and S in the objective functions to guarantee unique solutions. For example, Rajamani and Rawlings (2009) suggest penalizing the trace of Q to try to estimate a value of Q with the minimum possible rank. Note that it is not necessary to penalize P and $\mathbb{P}$ in such a manner, since P is uniquely determined by Q and $\mathbb{P}$ is uniquely determined by Q, R, and S (recall (2.19) and (2.27)).

Appendices

3.A Miscellaneous results

Lemma 3.1. *For any $n \times p$ matrix K, (A, C) is observable if and only if $(A - KC, C)$ is observable.*

Proof. Observe that

$$\begin{bmatrix} I_n & K \\ 0 & I_p \end{bmatrix} \begin{bmatrix} \lambda I_n - A \\ C \end{bmatrix} = \begin{bmatrix} \lambda I_n - (A - KC) \\ C \end{bmatrix}$$

The first matrix in this equation is invertible, which implies

$$\operatorname{rank} \begin{bmatrix} \lambda I_n - A \\ C \end{bmatrix} = \operatorname{rank} \begin{bmatrix} \lambda I_n - (A - KC) \\ C \end{bmatrix}$$

The result then follows from the Hautus lemma for observability (Hautus, 1972; Rawlings et al., 2017, Lemma 1.4), which states that the system (A, C) is observable if and only if

$$\operatorname{rank} \begin{bmatrix} \lambda I_n - A \\ C \end{bmatrix} = n \quad \forall\, \lambda \in \mathbb{C}$$ ■

Lemma 3.2. *If matrix A has LI columns and matrix B is invertible, then $(ABA^{\mathrm{T}})^{+} = (A^{+})^{\mathrm{T}} B^{-1} A^{+}$.*

Proof. We verify directly that the indicated inverse satisfies the four requirements of the Moore-Penrose inverse. Note that since A has LI columns, $A^+A = I$. First,

$$ABA^{\mathrm{T}}(A^+)^{\mathrm{T}}B^{-1}A^+ = ABB^{-1}A^+ = AA^+ = (AA^+)^{\mathrm{T}}$$
$$(A^+)^{\mathrm{T}}B^{-1}A^+ABA^{\mathrm{T}} = (A^+)^{\mathrm{T}}B^{-1}BA^{\mathrm{T}} = (A^+)^{\mathrm{T}}A^{\mathrm{T}} = (AA^+)^{\mathrm{T}} = AA^+$$

are both invertible matrices. Then

$$\begin{aligned} ABA^{\mathrm{T}}(A^+)^{\mathrm{T}}B^{-1}A^+ABA^{\mathrm{T}} &= AA^+ABA^{\mathrm{T}} = ABA^{\mathrm{T}} \\ (A^+)^{\mathrm{T}}B^{-1}A^+ABA^{\mathrm{T}}(A^+)^{\mathrm{T}}B^{-1}A^+ &= (A^+)^{\mathrm{T}}A^{\mathrm{T}}(A^+)^{\mathrm{T}}B^{-1}A^+ \\ &= (A^+AA^+)^{\mathrm{T}}B^{-1}A^+ = (A^+)^{\mathrm{T}}B^{-1}A^+ \end{aligned}$$ ■

3.B Analysis of Assumption 7

Observe that y_k and $\mathbb{z}_k$ can be expressed as

$$y_k = CA^k\mu_0 + \left(\sum_{j=0}^{k-1} CA^{k-j-1}Gw_j\right) + Hv_k$$
$$\mathbb{z}_k = C\mathbb{A}^k(x_0 - \mu_0) + \left(\sum_{j=0}^{k-1} C\mathbb{A}^{k-j-1}\mathbb{G}\bar{w}_j\right) + Hv_k$$

Because A and $\mathbb{A}$ are stable, as $k \to \infty$, $A^k \to 0$ and $\mathbb{A}^k \to 0$, and the influence of μ_0 (whether it is known accurately or not) on y_k and $\mathbb{z}_k$ diminishes exponentially with time, so as $k \to \infty$, the autocovariances of y_k and $\mathbb{z}_k$ approach the steady-state values shown in (2.21)–(2.24) and (2.28)–(2.31), respectively.

3.C Proof of Theorem 3.1

Proof. First we establish several useful identities. For any matrix $E \in \mathbb{R}^{k\times\ell}$,

$$\begin{aligned} (I_{k^2} + \mathscr{K}_k)(E \otimes E) &= (E \otimes E) + \mathscr{K}_k(E \otimes E) \\ &= (E \otimes E) + (E \otimes E)\mathscr{K}_\ell = (E \otimes E)(I_{\ell^2} + \mathscr{K}_\ell) \end{aligned}$$

Next,

$$\begin{aligned}\mathcal{K}_n F &= \mathcal{K}_n[I_{n^2} - (A \otimes A)]^{-1} = \mathcal{K}_n\{\mathcal{K}_n[I_{n^2} - (A \otimes A)]\mathcal{K}_n\}^{-1} \\ &= \mathcal{K}_n\mathcal{K}_n[I_{n^2} - (A \otimes A)]^{-1}\mathcal{K}_n = [I_{n^2} - (A \otimes A)]^{-1}\mathcal{K}_n = F\mathcal{K}_n\end{aligned}$$

Similarly, $\mathcal{K}_n\mathbb{F} = \mathbb{F}\mathcal{K}_n$. From Magnus and Neudecker (2019, Theorems 3.11 and 3.12, pages 55–56), for any natural number $k \geq 1$, we have

$$\begin{aligned}\mathcal{D}_k &= \mathcal{K}_k\mathcal{D}_k = \tfrac{1}{2}(I_{k^2} + \mathcal{K}_k)\mathcal{D}_k \\ I_{k^2} + \mathcal{K}_k &= 2\mathcal{D}_k\mathcal{D}_k^+ = \tfrac{1}{2}(I_{k^2} + \mathcal{K}_k)^2\end{aligned}$$

Now to the main results:

$$\begin{aligned}X_\Phi U_{\Phi Q} &= (I_p \otimes \Gamma)\mathcal{D}_p\mathcal{D}_p^+(C \otimes C)F(G \otimes G)\mathcal{D}_g \\ &= \tfrac{1}{2}(I_p \otimes \Gamma)(I_{p^2} + \mathcal{K}_p)(C \otimes C)F(G \otimes G)\mathcal{D}_g \\ &= \tfrac{1}{2}(I_p \otimes \Gamma)(C \otimes C)F(G \otimes G)(I_{g^2} + \mathcal{K}_g)\mathcal{D}_g \\ &= (I_p \otimes \Gamma)(C \otimes C)F(G \otimes G)\mathcal{D}_g \\ &= (C \otimes \Gamma C)F(G \otimes G)\mathcal{D}_g\end{aligned}$$

$$\begin{aligned}X_\Theta U_{\Theta Q} &= (I_p \otimes \mathcal{O}_{2,N})(C \otimes A)F(G \otimes G)\mathcal{D}_g \\ &= (C \otimes \mathcal{O}_{2,N}A)F(G \otimes G)\mathcal{D}_g\end{aligned}$$

$$\begin{aligned}X_\Phi U_{\Phi Q} + X_\Theta U_{\Theta Q} &= [(C \otimes \Gamma C) + (C \otimes \mathcal{O}_{2,N}A)]F(G \otimes G)\mathcal{D}_g \\ &= [C \otimes (\Gamma C + \mathcal{O}_{2,N}A)]F(G \otimes G)\mathcal{D}_g \\ &= (C \otimes \mathcal{O}_{1,N})F(G \otimes G)\mathcal{D}_g \\ &= X_Q\end{aligned}$$

$$\begin{aligned}X_\Phi U_{\Phi R} &= (I_p \otimes \Gamma)\mathcal{D}_p\mathcal{D}_p^+(H \otimes H)\mathcal{D}_h \\ &= \tfrac{1}{2}(I_p \otimes \Gamma)(I_{p^2} + \mathcal{K}_p)(H \otimes H)\mathcal{D}_h \\ &= \tfrac{1}{2}(I_p \otimes \Gamma)(H \otimes H)(I_{h^2} + \mathcal{K}_h)\mathcal{D}_h \\ &= (I_p \otimes \Gamma)(H \otimes H)\mathcal{D}_h \\ &= (H \otimes \Gamma H)\mathcal{D}_h \\ &= X_R\end{aligned}$$

$$\begin{aligned} X_\Theta U_{\Theta S} &= (I_p \otimes \mathcal{O}_{2,N})(H \otimes G) \\ &= (H \otimes \mathcal{O}_{2,N} G) \\ &= X_S \end{aligned}$$

$$\begin{aligned} X_{\Phi\Theta} U &= \mathscr{D}^+_{Np,p} \begin{bmatrix} X_\Phi & X_\Theta \end{bmatrix} \begin{bmatrix} U_{\Phi Q} & U_{\Phi R} & 0 \\ U_{\Theta Q} & 0 & U_{\Theta S} \end{bmatrix} \\ &= \mathscr{D}^+_{Np,p} \begin{bmatrix} X_\Phi U_{\Phi Q} + X_\Theta U_{\Theta Q} & X_\Phi U_{\Phi R} & X_\Theta U_{\Theta S} \end{bmatrix} \\ &= \mathscr{D}^+_{Np,p} \begin{bmatrix} X_Q & X_R & X_S \end{bmatrix} \\ &= X_{QRS} \end{aligned}$$

$$\begin{aligned} \mathbb{X}_\Phi \mathbb{U}_{\Phi Q} &= (I_p \otimes \Gamma) \mathscr{D}_p \mathscr{D}_p^+ (C \otimes C) \mathbb{F} (G \otimes G) \mathscr{D}_g \\ &= \tfrac{1}{2} (I_p \otimes \Gamma)(I_{p^2} + \mathscr{K}_p)(C \otimes C) \mathbb{F} (G \otimes G) \mathscr{D}_g \\ &= \tfrac{1}{2} (I_p \otimes \Gamma)(C \otimes C) \mathbb{F} (G \otimes G)(I_{g^2} + \mathscr{K}_g) \mathscr{D}_g \\ &= (I_p \otimes \Gamma)(C \otimes C) \mathbb{F} (G \otimes G) \mathscr{D}_g \\ &= (C \otimes \Gamma C) \mathbb{F} (G \otimes G) \mathscr{D}_g \end{aligned}$$

$$\begin{aligned} \mathbb{X}_\Theta \mathbb{U}_{\Theta Q} &= (I_p \otimes \mathbb{O}_{2,N})(C \otimes \mathbb{A}) \mathbb{F} (G \otimes G) \mathscr{D}_g \\ &= (C \otimes \mathbb{O}_{2,N} \mathbb{A}) \mathbb{F} (G \otimes G) \mathscr{D}_g \end{aligned}$$

$$\begin{aligned} \mathbb{X}_\Phi \mathbb{U}_{\Phi Q} + \mathbb{X}_\Theta \mathbb{U}_{\Theta Q} &= [(C \otimes \Gamma C) + (C \otimes \mathbb{O}_{2,N} \mathbb{A})] \mathbb{F} (G \otimes G) \mathscr{D}_g \\ &= [C \otimes (\Gamma C + \mathbb{O}_{2,N} \mathbb{A})] \mathbb{F} (G \otimes G) \mathscr{D}_g \\ &= (C \otimes \mathbb{O}_{1,N}) \mathbb{F} (G \otimes G) \mathscr{D}_g \\ &= \mathbb{X}_Q \end{aligned}$$

$$\begin{aligned} \mathbb{X}_\Phi \mathbb{U}_{\Phi R} &= (I_p \otimes \Gamma) \mathscr{D}_p \mathscr{D}_p^+ [(C \otimes C) \mathbb{F} (\mathcal{K} H \otimes \mathcal{K} H) + (H \otimes H)] \mathscr{D}_h \\ &= \tfrac{1}{2} (I_p \otimes \Gamma)(I_{p^2} + \mathscr{K}_p) [(C \otimes C) \mathbb{F} (\mathcal{K} H \otimes \mathcal{K} H) + (H \otimes H)] \mathscr{D}_h \\ &= \tfrac{1}{2} (I_p \otimes \Gamma) [(C \otimes C) \mathbb{F} (\mathcal{K} H \otimes \mathcal{K} H) + (H \otimes H)](I_{h^2} + \mathscr{K}_h) \mathscr{D}_h \\ &= (I_p \otimes \Gamma) [(C \otimes C) \mathbb{F} (\mathcal{K} H \otimes \mathcal{K} H) + (H \otimes H)] \mathscr{D}_h \\ &= [(C \otimes \Gamma C) \mathbb{F} (\mathcal{K} H \otimes \mathcal{K} H) + (H \otimes \Gamma H)] \mathscr{D}_h \end{aligned}$$

$$\begin{aligned}\mathbb{X}_{\Theta}\mathbb{U}_{\Theta R} &= (I_p \otimes \mathbb{O}_{2,N})[(C \otimes \mathbb{A})\mathbb{F}(\mathcal{K}H \otimes \mathcal{K}H) - (H \otimes \mathcal{K}H)]\mathscr{D}_h \\ &= [(C \otimes \mathbb{O}_{2,N}\mathbb{A})\mathbb{F}(\mathcal{K}H \otimes \mathcal{K}H) - (H \otimes \mathbb{O}_{2,N}\mathcal{K}H)]\mathscr{D}_h\end{aligned}$$

$$\begin{aligned}\mathbb{X}_{\Phi}\mathbb{U}_{\Phi R} + \mathbb{X}_{\Theta}\mathbb{U}_{\Theta R} &= \{[(C \otimes \Gamma C) + (C \otimes \mathbb{O}_{2,N}\mathbb{A})]\mathbb{F}(\mathcal{K}H \otimes \mathcal{K}H) \\ &\qquad + (H \otimes \Gamma H) - (H \otimes \mathbb{O}_{2,N}\mathcal{K}H)\}\mathscr{D}_h \\ &= \{[C \otimes (\Gamma C + \mathbb{O}_{2,N}\mathbb{A})]\mathbb{F}(\mathcal{K}H \otimes \mathcal{K}H) \\ &\qquad + [H \otimes (\Gamma - \mathbb{O}_{2,N}\mathcal{K})H]\}\mathscr{D}_h \\ &= [(C \otimes \mathbb{O}_{1,N})\mathbb{F}(\mathcal{K}H \otimes \mathcal{K}H) + (H \otimes \mathbb{\Gamma} H)]\mathscr{D}_h \\ &= \mathbb{X}_R\end{aligned}$$

$$\begin{aligned}\mathbb{X}_{\Phi}\mathbb{U}_{\Phi S} &= -(I_p \otimes \Gamma)\mathscr{D}_p\mathscr{D}_p^{+}(C \otimes C)\mathbb{F}(I_{n^2} + \mathscr{K}_n)(\mathcal{K}H \otimes G) \\ &= -\tfrac{1}{2}(I_p \otimes \Gamma)(I_{p^2} + \mathscr{K}_p)(C \otimes C)\mathbb{F}(I_{n^2} + \mathscr{K}_n)(\mathcal{K}H \otimes G) \\ &= -\tfrac{1}{2}(I_p \otimes \Gamma)(C \otimes C)\mathbb{F}(I_{n^2} + \mathscr{K}_n)^2(\mathcal{K}H \otimes G) \\ &= -(I_p \otimes \Gamma)(C \otimes C)\mathbb{F}(I_{n^2} + \mathscr{K}_n)(\mathcal{K}H \otimes G) \\ &= -(C \otimes \Gamma C)\mathbb{F}(I_{n^2} + \mathscr{K}_n)(\mathcal{K}H \otimes G)\end{aligned}$$

$$\begin{aligned}\mathbb{X}_{\Theta}\mathbb{U}_{\Theta S} &= (I_p \otimes \mathbb{O}_{2,N})[-(C \otimes \mathbb{A})\mathbb{F}(I_{n^2} + \mathscr{K}_n)(\mathcal{K}H \otimes G) + (H \otimes G)] \\ &= -(C \otimes \mathbb{O}_{2,N}\mathbb{A})\mathbb{F}(I_{n^2} + \mathscr{K}_n)(\mathcal{K}H \otimes G) + (H \otimes \mathbb{O}_{2,N}G)\end{aligned}$$

$$\begin{aligned}\mathbb{X}_{\Phi}\mathbb{U}_{\Phi S} + \mathbb{X}_{\Theta}\mathbb{U}_{\Theta S} &= -[(C \otimes \Gamma C) + (C \otimes \mathbb{O}_{2,N}\mathbb{A})]\mathbb{F}(I_{n^2} + \mathscr{K}_n)(\mathcal{K}H \otimes G) \\ &\qquad + (H \otimes \mathbb{O}_{2,N}G) \\ &= -[C \otimes (\Gamma C + \mathbb{O}_{2,N}\mathbb{A})]\mathbb{F}(I_{n^2} + \mathscr{K}_n)(\mathcal{K}H \otimes G) + (H \otimes \mathbb{O}_{2,N}G) \\ &= -(C \otimes \mathbb{O}_{1,N})\mathbb{F}(I_{n^2} + \mathscr{K}_n)(\mathcal{K}H \otimes G) + (H \otimes \mathbb{O}_{2,N}G) \\ &= \mathbb{X}_S\end{aligned}$$

$$\begin{aligned}\mathbb{X}_{\Phi\Theta}\mathbb{U} &= \mathscr{D}_{Np,p}^{+}\begin{bmatrix}\mathbb{X}_{\Phi} & \mathbb{X}_{\Theta}\end{bmatrix}\begin{bmatrix}\mathbb{U}_{\Phi Q} & \mathbb{U}_{\Phi R} & \mathbb{U}_{\Phi S} \\ \mathbb{U}_{\Theta Q} & \mathbb{U}_{\Theta R} & \mathbb{U}_{\Theta S}\end{bmatrix} \\ &= \mathscr{D}_{Np,p}^{+}\begin{bmatrix}\mathbb{X}_{\Phi}\mathbb{U}_{\Phi Q} + \mathbb{X}_{\Theta}\mathbb{U}_{\Theta Q} & \mathbb{X}_{\Phi}\mathbb{U}_{\Phi R} + \mathbb{X}_{\Theta}\mathbb{U}_{\Theta R} & \mathbb{X}_{\Phi}\mathbb{U}_{\Phi S} + \mathbb{X}_{\Theta}\mathbb{U}_{\Theta S}\end{bmatrix} \\ &= \mathscr{D}_{Np,p}^{+}\begin{bmatrix}\mathbb{X}_Q & \mathbb{X}_R & \mathbb{X}_S\end{bmatrix} \\ &= \mathbb{X}_{QRS}\end{aligned}$$ ■

3.D Proof of Theorem 3.2

Proof. Any matrix X has LI columns if and only if the zero vector is the only element of its null space. Therefore, the proof proceeds by characterizing the null spaces of X_{QRS}, $\mathbb{X}_{QRS}$, X_{QR}, $\mathbb{X}_{QR}$, $X_{\Phi\Theta}$, and $\mathbb{X}_{\Phi\Theta}$.

1. Consider

$$0 = X_{QRS}\begin{bmatrix} Q_{\mathrm{ss}} \\ R_{\mathrm{ss}} \\ S_{\mathrm{s}} \end{bmatrix} \qquad (3.17) \qquad\qquad 0 = \mathbb{X}_{QRS}\begin{bmatrix} \mathbb{Q}_{\mathrm{ss}} \\ \mathbb{R}_{\mathrm{ss}} \\ \mathbb{S}_{\mathrm{s}} \end{bmatrix} \qquad (3.18)$$

The unvectorized forms of (3.17) and (3.18) are respectively

$$0 = \mathcal{O}_{1,N} P C^{\mathrm{T}} + \Gamma R + \mathcal{O}_{2,N} S \qquad (3.19)$$

$$0 = \mathbb{O}_{1,N} \mathbb{P} C^{\mathrm{T}} + \mathbb{\Gamma} \mathbb{R} + \mathbb{O}_{2,N} \mathbb{S} \qquad (3.20)$$

where it is understood that P, $\mathbb{P}$, R, and $\mathbb{R}$ are symmetric matrices and that the bijective transformations $P \leftrightarrow Q$ and $\mathbb{P} \leftrightarrow \mathbb{Q} + \mathcal{K}\mathbb{R}\mathcal{K}^{\mathrm{T}} - \mathbb{S}\mathcal{K}^{\mathrm{T}} - \mathcal{K}\mathbb{S}^{\mathrm{T}}$ given in (2.19) and (2.27), respectively, have been used:

$$P = APA^{\mathrm{T}} + Q \qquad (3.21)$$

$$\begin{aligned} \mathbb{P} &= \mathbb{A}\mathbb{P}\mathbb{A}^{\mathrm{T}} + \mathbb{Q} + \mathcal{K}\mathbb{R}\mathcal{K}^{\mathrm{T}} - \mathbb{S}\mathcal{K}^{\mathrm{T}} - \mathcal{K}\mathbb{S}^{\mathrm{T}} \\ &= A\mathbb{P}A + \mathcal{K}(C\mathbb{P}C^{\mathrm{T}} + \mathbb{R})\mathcal{K}^{\mathrm{T}} + \mathbb{Q} \\ &\quad - \mathcal{K}(C\mathbb{P}A^{\mathrm{T}} + \mathbb{S}^{\mathrm{T}}) - (A\mathbb{P}C^{\mathrm{T}} + \mathbb{S})\mathcal{K}^{\mathrm{T}} \end{aligned} \qquad (3.22)$$

Equation (3.19) can be written as two separate equations

$$0 = CPC^{\mathrm{T}} + R \qquad (3.23) \qquad\qquad 0 = \mathcal{O}_{1,N-1}(APC^{\mathrm{T}} + S) \qquad (3.24)$$

as can (3.20)

$$0 = C\mathbb{P}C^{\mathrm{T}} + \mathbb{R} \qquad (3.25)$$

$$\begin{aligned} 0 &= \mathbb{O}_{1,N-1}(\mathbb{A}\mathbb{P}C^{\mathrm{T}} - \mathcal{K}\mathbb{R} + \mathbb{S}) \\ &= \mathbb{O}_{1,N-1}[A\mathbb{P}C^{\mathrm{T}} - \mathcal{K}(C\mathbb{P}C^{\mathrm{T}} + \mathbb{R}) + \mathbb{S}] \end{aligned} \qquad (3.26)$$

Let P and $\mathbb{P}$ be arbitrary nonzero symmetric matrices. Set $Q = P - APA^{\mathrm{T}}$, $R = -CPC^{\mathrm{T}}$, $S = -APC^{\mathrm{T}}$, $\mathbb{Q} = \mathbb{P} - A\mathbb{P}A^{\mathrm{T}}$, $\mathbb{R} = -C\mathbb{P}C^{\mathrm{T}}$, and $\mathbb{S} = -A\mathbb{P}C^{\mathrm{T}}$. Then (3.21)–(3.26) are satisfied. Next, $P \neq 0 \implies Q \neq 0$. Similarly, $\mathbb{P} \neq$

$0 \implies \mathbb{Q} + \mathcal{K}\mathbb{R}\mathcal{K}^{\mathrm{T}} - \mathbb{S}\mathcal{K}^{\mathrm{T}} - \mathcal{K}\mathbb{S}^{\mathrm{T}} \neq 0$. Therefore at least one of $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{S}$ must not be a zero matrix. Thus

$$\begin{bmatrix} Q_{\mathrm{ss}} \\ R_{\mathrm{ss}} \\ S_{\mathrm{s}} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \mathbb{Q}_{\mathrm{ss}} \\ \mathbb{R}_{\mathrm{ss}} \\ \mathbb{S}_{\mathrm{s}} \end{bmatrix}$$

are nonzero elements of the null spaces of X_{QRS} and $\mathbb{X}_{QRS}$, respectively.

2. Consider

$$0 = X_{QR} \begin{bmatrix} Q_{\mathrm{ss}} \\ R_{\mathrm{ss}} \end{bmatrix} \qquad (3.27) \qquad\qquad 0 = \mathbb{X}_{QR} \begin{bmatrix} \mathbb{Q}_{\mathrm{ss}} \\ \mathbb{R}_{\mathrm{ss}} \end{bmatrix} \qquad (3.28)$$

The unvectorized forms of (3.27) and (3.28) are respectively

$$0 = \mathcal{O}_{1,N} P C^{\mathrm{T}} + \Gamma R \qquad (3.29) \qquad\qquad 0 = \mathbb{O}_{1,N} \mathbb{P} C^{\mathrm{T}} + \mathbb{\Gamma} \mathbb{R} \qquad (3.30)$$

where it is understood that P, $\mathbb{P}$, R, and $\mathbb{R}$ are symmetric matrices and that the bijective transformations $P \leftrightarrow Q$ and $\mathbb{P} \leftrightarrow \mathbb{Q} + \mathcal{K}\mathbb{R}\mathcal{K}^{\mathrm{T}}$ given in (2.19) and (2.27), respectively, have been used:

$$P = APA^{\mathrm{T}} + Q \qquad (3.31)$$

$$\begin{aligned} \mathbb{P} &= \mathbb{A}\mathbb{P}\mathbb{A}^{\mathrm{T}} + \mathbb{Q} + \mathcal{K}\mathbb{R}\mathcal{K}^{\mathrm{T}} \\ &= A\mathbb{P}A^{\mathrm{T}} + \mathcal{K}(C\mathbb{P}C^{\mathrm{T}} + \mathbb{R})\mathcal{K}^{\mathrm{T}} + \mathbb{Q} - \mathcal{K}C\mathbb{P}A^{\mathrm{T}} - A\mathbb{P}C^{\mathrm{T}}\mathcal{K}^{\mathrm{T}} \end{aligned} \qquad (3.32)$$

Equation (3.29) can be written as two separate equations

$$0 = CPC^{\mathrm{T}} + R \qquad (3.33) \qquad\qquad 0 = \mathcal{O}_{1,N-1} A P C^{\mathrm{T}} \qquad (3.34)$$

as can (3.30)

$$0 = C\mathbb{P}C^{\mathrm{T}} + \mathbb{R} \qquad (3.35)$$

$$0 = \mathbb{O}_{1,N-1}(\mathbb{A}\mathbb{P}C^{\mathrm{T}} - \mathcal{K}\mathbb{R}) = \mathbb{O}_{1,N-1}[A\mathbb{P}C^{\mathrm{T}} - \mathcal{K}(C\mathbb{P}C^{\mathrm{T}} + \mathbb{R})] \qquad (3.36)$$

Note that substituting (3.35) into (3.36) gives

$$0 = \mathbb{O}_{1,N-1} A \mathbb{P} C^{\mathrm{T}} \qquad (3.37)$$

First assume that C has LI columns and A is invertible. The former implies that $\mathcal{O}_{1,N-1}$ and $\mathbb{O}_{1,N-1}$ have LI columns. Therefore (3.34) and (3.37) imply

$P = 0$ and $\mathbb{P} = 0$, respectively. Then (3.33) and (3.35) imply $R = 0$ and $\mathbb{R} = 0$, respectively. Finally (3.31) and (3.32) then imply $Q = 0$ and $\mathbb{Q} = 0$, respectively. Thus the zero vector is the only element in the null spaces of both X_{QR} and $\mathbb{X}_{QR}$.

Now assume A is not invertible. Then there exists a nonzero vector z such that $Az = 0$. Set $P = \mathbb{P} = zz^{\mathrm{T}} \neq 0$, $Q = P$, $R = -CPC^{\mathrm{T}}$, $\mathbb{Q} = \mathbb{P}$, and $\mathbb{R} = -C\mathbb{P}C^{\mathrm{T}}$. Then (3.31)–(3.37) are satisfied. Thus

$$\begin{bmatrix} Q_{\mathrm{ss}} \\ R_{\mathrm{ss}} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \mathbb{Q}_{\mathrm{ss}} \\ \mathbb{R}_{\mathrm{ss}} \end{bmatrix}$$

are nonzero elements of the null spaces of X_{QR} and $\mathbb{X}_{QR}$, respectively.

Now assume that C does not have LI columns. Then there exists a nonzero vector z such that $Cz = 0$. Set $P = \mathbb{P} = zz^{\mathrm{T}} \neq 0$, $Q = P - APA^{\mathrm{T}}$, $R = 0$, $\mathbb{Q} = \mathbb{P} - A\mathbb{P}A^{\mathrm{T}}$, and $\mathbb{R} = 0$. Then (3.31)–(3.37) are satisfied. Next, $P \neq 0 \implies Q \neq 0$. Similarly, $\mathbb{P} \neq 0 \implies \mathbb{Q} + \mathcal{K}\mathbb{R}\mathcal{K}^{\mathrm{T}} \neq 0$. We know $\mathbb{R} = 0$, so we conclude that $\mathbb{Q} \neq 0$. Thus

$$\begin{bmatrix} Q_{\mathrm{ss}} \\ R_{\mathrm{ss}} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \mathbb{Q}_{\mathrm{ss}} \\ \mathbb{R}_{\mathrm{ss}} \end{bmatrix}$$

are nonzero elements of the null spaces of X_{QR} and $\mathbb{X}_{QR}$, respectively.

3. Consider

$$0 = X_{\Phi\Theta} \begin{bmatrix} \Phi_{\mathrm{ss}} \\ \Theta_{\mathrm{s}} \end{bmatrix} \tag{3.38}$$

The unvectorized form of (3.38) is

$$0 = \Gamma\Phi + \mathcal{O}_{2,N}\Theta \tag{3.39}$$

where it is understood that Φ is a symmetric matrix. Equation (3.39) can be written as two separate equations:

$$0 = \Phi \tag{3.40}$$

$$0 = \mathcal{O}_{1,N-1}\Theta \tag{3.41}$$

Assume (A, C) is observable. Then, since $N \geq n + 1$, $\mathcal{O}_{1,N-1}$ has LI columns and therefore (3.41) implies $\Theta = 0$. Thus, in light of (3.40), the zero vector is the only element of the null space of $X_{\Phi\Theta}$.

Now assume (A, C) is not observable. Then $\mathcal{O}_{1,N-1}$ does not have LI columns for any N, and therefore there exists a nonzero Θ that satisfies (3.41). Then for any such Θ and $\Phi = 0$,

$$\begin{bmatrix} \Phi_{\mathrm{ss}} \\ \Theta_{\mathrm{s}} \end{bmatrix}$$

is a nonzero element of the null space of $X_{\Phi\Theta}$.

$X_{\Phi\Theta}$ and $\mathbb{X}_{\Phi\Theta}$ are structurally identical; the only difference is that $\mathbb{A}$ is used in place of A in forming the latter. Thus the result is the same for $\mathbb{X}_{\Phi\Theta}$, because (A, C) is observable if and only if $(\mathbb{A}, C)$ is observable (Lemma 3.1). ■

3.E Proof of Theorem 3.3

Proof. Per Theorem A.1, the unique solution to (3.9) is

$$\begin{bmatrix} \hat{\Phi}_{\mathrm{ss}} \\ \hat{\Theta}_{\mathrm{s}} \end{bmatrix} = (X_{\Phi\Theta}^{\mathrm{T}} W_{\Phi\Theta} X_{\Phi\Theta})^{-1} X_{\Phi\Theta}^{\mathrm{T}} W_{\Phi\Theta} \hat{b}$$

Lemmas A.1–A.3 imply that the indicated inverse exists:

$$\operatorname{col} W_{\Phi\Theta} = \operatorname{col} W_{\Phi\Theta}^{+} = \operatorname{col}(V + X_{\Phi\Theta} X_{\Phi\Theta}^{\mathrm{T}}) \supseteq \operatorname{col} X_{\Phi\Theta} X_{\Phi\Theta}^{\mathrm{T}} = \operatorname{col} X_{\Phi\Theta}$$
$$\implies X_{\Phi\Theta}^{\mathrm{T}} W_{\Phi\Theta} X_{\Phi\Theta} > 0$$

Using Lemmas A.1–A.3 again along with $X_{QRS} = X_{\Phi\Theta} U$ (Theorem 3.1) and the fact that U has LI rows implies that the matrix $X_{\Phi\Theta}^{\mathrm{T}} W_{QRS} X_{\Phi\Theta}$ is also invertible:

$$\operatorname{col} W_{QRS} = \operatorname{col} W_{QRS}^{+} = \operatorname{col}(V + X_{QRS} X_{QRS}^{\mathrm{T}}) = \operatorname{col}(V + X_{\Phi\Theta} U U^{\mathrm{T}} X_{\Phi\Theta}^{\mathrm{T}})$$
$$\supseteq \operatorname{col} X_{\Phi\Theta} U U^{\mathrm{T}} X_{\Phi\Theta}^{\mathrm{T}} = \operatorname{col} X_{\Phi\Theta} U = \operatorname{col} X_{\Phi\Theta} \implies X_{\Phi\Theta}^{\mathrm{T}} W_{QRS} X_{\Phi\Theta} > 0$$

Then, again using Theorem A.1, the general solution to (3.7) is

$$\begin{aligned} \begin{bmatrix} \hat{Q}_{\mathrm{ss}} \\ \hat{R}_{\mathrm{ss}} \\ \hat{S}_{\mathrm{s}} \end{bmatrix} &= (X_{QRS}^{\mathrm{T}} W_{QRS} X_{QRS})^{+} X_{QRS}^{\mathrm{T}} W_{QRS} \hat{b} \\ &\qquad + [I - (X_{QRS}^{\mathrm{T}} W_{QRS} X_{QRS})^{+} X_{QRS}^{\mathrm{T}} W_{QRS} X_{QRS}] q \\ &= (U^{\mathrm{T}} X_{\Phi\Theta}^{\mathrm{T}} W_{QRS} X_{\Phi\Theta} U)^{+} U^{\mathrm{T}} X_{\Phi\Theta}^{\mathrm{T}} W_{QRS} \hat{b} \\ &\qquad + [I - (U^{\mathrm{T}} X_{\Phi\Theta}^{\mathrm{T}} W_{QRS} X_{\Phi\Theta} U)^{+} U^{\mathrm{T}} X_{\Phi\Theta}^{\mathrm{T}} W_{QRS} X_{\Phi\Theta} U] q \end{aligned}$$

$$
\begin{aligned}
&= U^{+}(X_{\Phi\Theta}^{\mathrm{T}} W_{QRS} X_{\Phi\Theta})^{-1}(U^{\mathrm{T}})^{+} U^{\mathrm{T}} X_{\Phi\Theta}^{\mathrm{T}} W_{QRS} \hat{b} \\
&\qquad + [I - U^{+}(X_{\Phi\Theta}^{\mathrm{T}} W_{QRS} X_{\Phi\Theta})^{-1}(U^{\mathrm{T}})^{+} U^{\mathrm{T}} X_{\Phi\Theta}^{\mathrm{T}} W_{QRS} X_{\Phi\Theta} U] q \\
&= U^{+}(X_{\Phi\Theta}^{\mathrm{T}} W_{QRS} X_{\Phi\Theta})^{-1} X_{\Phi\Theta}^{\mathrm{T}} W_{QRS} \hat{b} \\
&\qquad + [I - U^{+}(X_{\Phi\Theta}^{\mathrm{T}} W_{QRS} X_{\Phi\Theta})^{-1} X_{\Phi\Theta}^{\mathrm{T}} W_{QRS} X_{\Phi\Theta} U] q \\
&= U^{+}(X_{\Phi\Theta}^{\mathrm{T}} W_{QRS} X_{\Phi\Theta})^{-1} X_{\Phi\Theta}^{\mathrm{T}} W_{QRS} \hat{b} + (I - U^{+}U) q
\end{aligned}
$$

where q is an arbitrary vector of the appropriate dimension. In the above, we have made use of Lemma 3.2 and the fact that U having LI rows implies $UU^{+} = (U^{\mathrm{T}})^{+}U^{\mathrm{T}} = I$. Finally,

$$
\begin{aligned}
\begin{bmatrix} [f_{\Phi}(\hat{Q},\hat{R},\hat{S})]_{\mathrm{ss}} \\ [f_{\Theta}(\hat{Q},\hat{R},\hat{S})]_{\mathrm{s}} \end{bmatrix} &= U \begin{bmatrix} \hat{Q}_{\mathrm{ss}} \\ \hat{R}_{\mathrm{ss}} \\ \hat{S}_{\mathrm{s}} \end{bmatrix} \\
&= UU^{+}(X_{\Phi\Theta}^{\mathrm{T}} W_{QRS} X_{\Phi\Theta})^{-1} X_{\Phi\Theta}^{\mathrm{T}} W_{QRS} \hat{b} + U(I - U^{+}U) q \\
&= (X_{\Phi\Theta}^{\mathrm{T}} W_{QRS} X_{\Phi\Theta})^{-1} X_{\Phi\Theta}^{\mathrm{T}} W_{QRS} \hat{b} + (U - UU^{+}U) q \\
&= (X_{\Phi\Theta}^{\mathrm{T}} W_{QRS} X_{\Phi\Theta})^{-1} X_{\Phi\Theta}^{\mathrm{T}} W_{QRS} \hat{b} + (U - U) q \\
&= (X_{\Phi\Theta}^{\mathrm{T}} W_{QRS} X_{\Phi\Theta})^{-1} X_{\Phi\Theta}^{\mathrm{T}} W_{QRS} \hat{b} \\
&= [X_{\Phi\Theta}^{\mathrm{T}}(V + X_{QRS} X_{QRS}^{\mathrm{T}})^{+} X_{\Phi\Theta}]^{-1} X_{\Phi\Theta}^{\mathrm{T}}(V + X_{QRS} X_{QRS}^{\mathrm{T}})^{+} \hat{b} \\
&= [X_{\Phi\Theta}^{\mathrm{T}}(V + X_{\Phi\Theta} U U^{\mathrm{T}} X_{\Phi\Theta}^{\mathrm{T}})^{+} X_{\Phi\Theta}]^{-1} X_{\Phi\Theta}^{\mathrm{T}}(V + X_{\Phi\Theta} U U^{\mathrm{T}} X_{\Phi\Theta}^{\mathrm{T}})^{+} \hat{b} \\
&= [X_{\Phi\Theta}^{\mathrm{T}}(V + X_{\Phi\Theta} X_{\Phi\Theta}^{\mathrm{T}})^{+} X_{\Phi\Theta}]^{-1} X_{\Phi\Theta}^{\mathrm{T}}(V + X_{\Phi\Theta} X_{\Phi\Theta}^{\mathrm{T}})^{+} \hat{b} \\
&= (X_{\Phi\Theta}^{\mathrm{T}} W_{\Phi\Theta} X_{\Phi\Theta})^{-1} X_{\Phi\Theta}^{\mathrm{T}} W_{\Phi\Theta} \hat{b} \\
&= \begin{bmatrix} \hat{\Phi}_{\mathrm{ss}} \\ \hat{\Theta}_{\mathrm{s}} \end{bmatrix}
\end{aligned}
$$

where we have used Theorem A.3 (U having LI rows implies $UU^{\mathrm{T}} > 0$, so UU^{T} corresponds to E in Theorem A.3). ■

Perfection is not attainable, but if we chase perfection we can catch excellence.

Vince Lombardi

CHAPTER 4

TRACTABLE CALCULATION OF THE OPTIMAL ALS WEIGHTING MATRIX

This chapter serves Objective 2 of the dissertation, which is to develop the first generally tractable method to calculate the optimal weighting matrix for the ALS problem. In Chapters 2 and 3, results are presented in two parallel tracks, for the process system without inputs and A stable, (1.3) and (1.4), and also for the $\mathcal{K}$-error system, (2.10) and (2.11). Although the results are parallel for the two systems, both are explicitly presented because there are subtle differences in the derivations and proofs between the two. In contrast, for this chapter, the results are entirely analogous for the two systems, including the derivations and proofs. The only difference is the names of the variables, e.g., $\mathbb{\Lambda}_j$ replaces Λ_j when considering the $\mathcal{K}$-error system. As such, here we do away with the $\mathcal{K}$-error system track and consider only the process system without inputs and A stable.

The chapter is organized as follows:

- Section 4.1 gives relevant background information and discusses assumptions for the chapter.
- Section 4.2 summarizes previous work on calculation of the optimal ALS weighting matrix.
- Section 4.3 derives a novel tractable method to calculate the optimal ALS weighting matrix.

The chapter also contains one appendix: 4.A.

4.1 Background and assumptions

For this chapter, assume the outputs $y_{0:T_f}$ have the steady-state autocovariance given in (2.21)–(2.24). Per (3.7)–(3.12) and Corollary A.2, the optimal ALS weighting matrix is $W_\Omega = (P_{\hat{b}} + X_\Omega X_\Omega^{\mathrm{T}})^+$, where $P_{\hat{b}} := \mathrm{cov}(\hat{b})$ and Ω refers to any of (Q, R, S), $(Q, R, 0)$, or (Φ, Θ). Therefore, since X_Ω is known, estimating W_Ω is akin to estimating $P_{\hat{b}}$.

Recall from Chapter 3 that the ALS method does not require assumption of a particular probability distribution for the noises (w_k) and (v_k). In this chapter, we show that $P_{\hat{b}}$ depends on the fourth moment of the outputs, which at this point we can say nothing about since we have heretofore not assumed a specific probability distribution for the noises. Hence, in this chapter, we assume that (w_k) and (v_k) are jointly normally distributed, i.e.

$$\begin{bmatrix} w_k \\ v_\ell \\ w_i \\ v_j \end{bmatrix} \sim \mathrm{N}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} Q & S\delta_{k\ell} & Q\delta_{ki} & S\delta_{kj} \\ S^{\mathrm{T}}\delta_{k\ell} & R & S^{\mathrm{T}}\delta_{\ell i} & R\delta_{\ell j} \\ Q\delta_{ki} & S\delta_{\ell i} & Q & S\delta_{ij} \\ S^{\mathrm{T}}\delta_{kj} & R\delta_{\ell j} & S^{\mathrm{T}}\delta_{ij} & R \end{bmatrix} \right) \tag{4.1}$$

where δ_{ki} is the Kronecker delta. Note that under the assumption of jointly normal (w_k) and (v_k), the output sequence (y_k) is Gaussian, i.e., normally distributed. Hence, per Remark 0.4, $y_{0:T_f}$ is stationary.

Recall that ALS is based on the linear model (3.6):

$$\begin{gathered} \hat{b} = X_\Omega \Omega + e \\ \mathrm{E}[e] = 0 \qquad \mathrm{E}[ee^{\mathrm{T}}] = \mathrm{cov}(\hat{b}) = P_{\hat{b}} \end{gathered} \tag{4.2}$$

Consider for a moment the general linear model from Appendix A:

$$\begin{gathered} y = X\beta + \varepsilon \\ \mathrm{E}[\varepsilon] = 0 \qquad \mathrm{E}[\varepsilon\varepsilon^{\mathrm{T}}] = V \end{gathered} \tag{4.3}$$

Our overall focus of this dissertation is estimation of Ω in (4.2), which corresponds to β in (4.3). In this chapter, the focus is on $P_{\hat{b}}$, which corresponds to V in (4.3). In applications with linear models, V is often partially or completely unknown. Statisticians have developed a bevy of techniques to tackle the classical problem of

simultaneous estimation of β and V; see Magnus and Neudecker (2019, Chapter 14) and the references therein as a primer on this topic. Regarding (4.2), in this chapter we demonstrate that $P_{\hat{b}}$ is a function of $\Lambda_{0:T_f}$. Recall from (2.21)–(2.24) that $\Lambda_{0:T_f}$ itself is a function of Ω. That is, $P_{\hat{b}} = f_{P_{\hat{b}}}(\Omega)$ for a known function $f_{P_{\hat{b}}}$ which is elucidated in the sequel. The corresponding scenario in (4.3) is $V = f(\beta)$ for some known function f. This case, where the covariance matrix V is a deterministic function of the regressor β, is special. We are not aware of any literature concerning estimation of β and V in this scenario for arbitrary f.

4.2 Previous work

Rajamani and Rawlings (2009) were the first to give explicitly an expression for the optimal ALS weighting matrix. However, Remark 12 from that paper summarizes the associated difficulties: "The computation of [the optimal ALS weighting matrix] becomes prohibitively large even for a small dimensional problem with large data sets. This [issue] is a drawback for any practical application until efficient means for the computation are found. Although the weight may be estimated from data, a large data set is required before getting reliable estimates for the weight." Next, we derive an expression for $P_{\hat{b}}$ in the same manner as in Rajamani and Rawlings (2009) in order to demonstrate said difficulties.

Define

$$N_d \coloneqq T_f + 1 \qquad \tilde{p} \coloneqq Np \qquad \tilde{n} \coloneqq T_f - N + 1 = N_d - N$$

$$Y \coloneqq \begin{bmatrix} y_0 & y_1 & y_2 & \cdots & y_{\tilde{n}} & y_{\tilde{n}+1} & y_{\tilde{n}+2} & \cdots & y_{T_f-2} & y_{T_f-1} & y_{T_f} \\ y_1 & y_2 & y_3 & \cdots & y_{\tilde{n}+1} & y_{\tilde{n}+2} & y_{\tilde{n}+3} & \cdots & y_{T_f-1} & y_{T_f} & 0 \\ y_2 & y_3 & y_4 & \cdots & y_{\tilde{n}+2} & y_{\tilde{n}+3} & y_{\tilde{n}+4} & \cdots & y_{T_f} & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ y_{N-3} & y_{N-2} & y_{N-1} & \cdots & y_{T_f-2} & y_{T_f-1} & y_{T_f} & \cdots & 0 & 0 & 0 \\ y_{N-2} & y_{N-1} & y_N & \cdots & y_{T_f-1} & y_{T_f} & 0 & \cdots & 0 & 0 & 0 \\ y_{N-1} & y_N & y_{N+1} & \cdots & y_{T_f} & 0 & 0 & \cdots & 0 & 0 & 0 \end{bmatrix}$$

We partition Y columnwise. Let c_i denote the $(i+1)$-th column of Y:

$$Y = \begin{bmatrix} c_0 & c_1 & \cdots & c_{T_f} \end{bmatrix}$$

For $i \geq \tilde{n}$, the nonzero elements of c_i are copied from $c_{\tilde{n}}$, which can be expressed as

$$c_{\tilde{n}+j} = \mathcal{E}_j c_{\tilde{n}}, \quad 0 \leq j \leq N-1$$

where $\mathcal{E}_j$ is the appropriate $\tilde{p} \times \tilde{p}$ matrix of zeros and ones. Specifically, $\mathcal{E}_j$ is all zeros except for ones along the jp-th diagonal above the main diagonal. Note that $\mathcal{E}_0 = I_{\tilde{p}}$. The matrix $\hat{\Psi}$ may be expressed as a function of Y:

$$\hat{\Psi} = \mathcal{F} Y Y^{\mathrm{T}} \mathcal{G}^{\mathrm{T}}$$

$$\mathcal{F} \coloneqq \begin{bmatrix} \frac{1}{T_f+1} & & & \\ & \frac{1}{T_f} & & \\ & & \ddots & \\ & & & \frac{1}{\tilde{n}+1} \end{bmatrix} \otimes I_p \qquad \mathcal{G} \coloneqq \begin{bmatrix} I_p & 0_{p,p(N-1)} \end{bmatrix}$$

The vector $\hat{b}$ may be expressed as

$$\hat{b} = \hat{\Psi}_{\mathrm{ss}} = \mathscr{D}^{+}_{\tilde{p},p} \hat{\Psi}_{\mathrm{s}} = \mathcal{H} a$$

$$\mathcal{H} \coloneqq \mathscr{D}^{+}_{\tilde{p},p} (\mathcal{G} \otimes \mathcal{F}) \qquad a \coloneqq (Y Y^{\mathrm{T}})_{\mathrm{s}}$$

Then

$$P_{\hat{b}} = \mathcal{H} P_a \mathcal{H}^{\mathrm{T}} \tag{4.4}$$

$$\begin{aligned} P_a &\coloneqq \operatorname{cov}(a) \\ &= \operatorname{cov}[(Y Y^{\mathrm{T}})_{\mathrm{s}}] \\ &= \operatorname{cov}[(Y I_{N_d} Y^{\mathrm{T}})_{\mathrm{s}}] \\ &= \operatorname{cov}\{[(Y^{\mathrm{T}})^{\mathrm{T}} I_{N_d} (Y^{\mathrm{T}})]_{\mathrm{s}}\} \\ &= \mathcal{L} (I_{N_d^2 \tilde{p}^2} + \mathscr{K}_{N_d \tilde{p}}) (\mathscr{K}_{\tilde{p},N_d} P_{Y_{\mathrm{s}}} \mathscr{K}_{N_d,\tilde{p}} \otimes \mathscr{K}_{\tilde{p},N_d} P_{Y_{\mathrm{s}}} \mathscr{K}_{N_d,\tilde{p}}) \mathcal{L}^{\mathrm{T}} \end{aligned} \tag{4.5}$$

$$\mathcal{L} \coloneqq [I_{\tilde{p}^2} \otimes (I_{N_d})_{\mathrm{s}}]^{\mathrm{T}} (I_{\tilde{p}} \otimes \mathscr{K}_{\tilde{p},N_d} \otimes I_{N_d})$$

$$P_{Y_{\mathrm{s}}} \coloneqq \operatorname{cov}(Y_{\mathrm{s}})$$

To arrive at (4.5), we have used Corollary 10.2 from Ghazal and Neudecker (2000) and the fact that $\operatorname{cov}[(Y^{\mathrm{T}})_{\mathrm{s}}] = \operatorname{cov}(\mathscr{K}_{\tilde{p},N_d} Y_{\mathrm{s}}) = \mathscr{K}_{\tilde{p},N_d} P_{Y_{\mathrm{s}}} \mathscr{K}_{N_d,\tilde{p}}$. The vector Y_{s} may be expressed as

$$Y_{\mathrm{s}} = \mathit{\Pi} \tilde{Y}_{\mathrm{s}} \qquad \mathit{\Pi} \coloneqq \begin{bmatrix} I_{\tilde{p}} & & & & \\ & I_{\tilde{p}} & & & \\ & & \ddots & & \\ & & & I_{\tilde{p}} & \\ & & & & \mathcal{E}_0 \\ & & & & \mathcal{E}_1 \\ & & & & \vdots \\ & & & & \mathcal{E}_{N-1} \end{bmatrix} \qquad \tilde{Y} \coloneqq \begin{bmatrix} c_0 & c_1 & \cdots & c_{\tilde{n}} \end{bmatrix}$$

Moreover, $\tilde{Y}_\mathrm{s}$ may be expressed as

$$\tilde{Y}_\mathrm{s} = \mathcal{T}\tilde{y} \qquad \mathcal{T} \coloneqq \begin{bmatrix} \begin{bmatrix} I_p & & \\ & \ddots & \\ & & I_p \end{bmatrix} & & & \\ & \begin{bmatrix} I_p & & \\ & \ddots & \\ & & I_p \end{bmatrix} & & \\ & & \vdots & \ddots \\ & & \ddots & \vdots \\ & & & \begin{bmatrix} I_p & & \\ & \ddots & \\ & & I_p \end{bmatrix} \end{bmatrix} \qquad \tilde{y} \coloneqq \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{T_f} \end{bmatrix} \tag{4.6}$$

Combining these equations gives $Y_\mathrm{s} = \mathcal{N}\mathcal{T}\tilde{y}$. Therefore

$$P_{Y_\mathrm{s}} = \mathcal{N}\mathcal{T}P_{\tilde{y}}\mathcal{T}^\mathrm{T}\mathcal{N}^\mathrm{T}$$

$$P_{\tilde{y}} \coloneqq \operatorname{cov}(\tilde{y}) = \mathrm{E}\left[\begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{T_f} \end{bmatrix} \begin{bmatrix} y_0^\mathrm{T} & y_1^\mathrm{T} & \cdots & y_{T_f}^\mathrm{T} \end{bmatrix}\right] = \begin{bmatrix} \Lambda_0 & \Lambda_1^\mathrm{T} & \cdots & \Lambda_{T_f}^\mathrm{T} \\ \Lambda_1 & \Lambda_0 & \cdots & \Lambda_{T_f-1}^\mathrm{T} \\ \vdots & \vdots & \ddots & \vdots \\ \Lambda_{T_f} & \Lambda_{T_f-1} & \cdots & \Lambda_0 \end{bmatrix} \tag{4.7}$$

Calculating P_a using (4.5) requires calculation of

$$\mathscr{K}_{\tilde{p},N_d} P_{Y_\mathrm{s}} \mathscr{K}_{N_d,\tilde{p}} \otimes \mathscr{K}_{\tilde{p},N_d} P_{Y_\mathrm{s}} \mathscr{K}_{N_d,\tilde{p}} \tag{4.8}$$

which is a dense matrix with dimensions $(N_d\tilde{p})^2 \times (N_d\tilde{p})^2$. The variable $\tilde{p}$ is generally of moderate size, but N_d is often several orders of magnitude larger than $\tilde{p}$ since it is desirable to use a large data set to get the best possible parameter estimates. As N_d grows, the memory required to store (4.8) very quickly balloons to infeasible levels, even if N and p are small.

The first tractable technique for estimating the weight from data was offered in Zagrobelny and Rawlings (2014) (conference version: Zagrobelny and Rawlings (2015a)). The numerical examples in that paper demonstrate a noticeably reduced variance for the covariance estimates obtained with the estimated ALS weighting matrix relative to using an identity matrix as the weighting. However, in Section 6.1.6

of this dissertation, we provide a detailed analysis and critique of Zagrobelny and Rawlings' technique, which reveals that it tacitly relies on assumptions which generally do not hold in practice. This observation begs the question of whether further improvement is possible, which is what motivated all of the work in this dissertation related to estimating the optimal ALS weighting matrix from data.

4.3 Tractable expressions for $P_{\hat{b}}$

Our next objective is to derive an expression for $P_{\hat{b}}$ with memory storage requirements independent of N_d. First, YY^{T} can be expressed as

$$\begin{aligned} YY^{\mathrm{T}} &= c_0 c_0^{\mathrm{T}} + c_1 c_1^{\mathrm{T}} + \cdots + c_{T_f} c_{T_f}^{\mathrm{T}} && (4.9)\\ &= c_0 c_0^{\mathrm{T}} + c_1 c_1^{\mathrm{T}} + \cdots + c_{\tilde{n}-1} c_{\tilde{n}-1}^{\mathrm{T}} \\ &\quad + \mathcal{E}_0 c_{\tilde{n}} c_{\tilde{n}}^{\mathrm{T}} \mathcal{E}_0^{\mathrm{T}} + \mathcal{E}_1 c_{\tilde{n}} c_{\tilde{n}}^{\mathrm{T}} \mathcal{E}_1^{\mathrm{T}} + \cdots + \mathcal{E}_{N-1} c_{\tilde{n}} c_{\tilde{n}}^{\mathrm{T}} \mathcal{E}_{N-1}^{\mathrm{T}} && (4.10) \end{aligned}$$

Vectorizing (4.10) yields

$$a = (YY^{\mathrm{T}})_{\mathrm{s}} = (c_0 \otimes c_0) + (c_1 \otimes c_1) + \cdots + (c_{\tilde{n}-1} \otimes c_{\tilde{n}-1}) + \tilde{\mathcal{E}}(c_{\tilde{n}} \otimes c_{\tilde{n}}) = \mathcal{U} c_{\otimes}$$

$$\tilde{\mathcal{E}} \coloneqq \sum_{j=0}^{N-1} (\mathcal{E}_j \otimes \mathcal{E}_j) \qquad \mathcal{U} \coloneqq \begin{bmatrix} (\imath_{\tilde{n}}^{\mathrm{T}} \otimes I_{\tilde{p}^2}) & \tilde{\mathcal{E}} \end{bmatrix} \qquad c_{\otimes} \coloneqq \begin{bmatrix} c_0 \otimes c_0 \\ c_1 \otimes c_1 \\ \vdots \\ c_{\tilde{n}} \otimes c_{\tilde{n}} \end{bmatrix}$$

The covariance of $c_{\otimes}$ is

$$P_{c_{\otimes}} \coloneqq \operatorname{cov}(c_{\otimes}) = \begin{bmatrix} P_{\otimes,0,0} & P_{\otimes,1,0}^{\mathrm{T}} & \cdots & P_{\otimes,\tilde{n},0}^{\mathrm{T}} \\ P_{\otimes,1,0} & P_{\otimes,0,1} & \cdots & P_{\otimes,\tilde{n}-1,1}^{\mathrm{T}} \\ \vdots & \vdots & \ddots & \vdots \\ P_{\otimes,\tilde{n},0} & P_{\otimes,\tilde{n}-1,1} & \cdots & P_{\otimes,0,\tilde{n}} \end{bmatrix}$$

$$P_{\otimes,j,k} \coloneqq \operatorname{cov}(c_{k+j} \otimes c_{k+j}, c_k \otimes c_k) = (I_{\tilde{p}^2} + \mathscr{K}_{\tilde{p}})(P_{c,j,k} \otimes P_{c,j,k}) \qquad (4.11)$$

$$P_{c,j,k} \coloneqq \operatorname{cov}(c_{k+j}, c_k) = \begin{bmatrix} \Lambda_j & \Lambda_{j-1} & \cdots & \Lambda_{j-N+q_k+1} & 0_{p,q_k p} \\ \Lambda_{j+1} & \Lambda_j & \cdots & \Lambda_{j-N+q_k+2} & 0_{p,q_k p} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \Lambda_{j+N-q_{k+j}-1} & \Lambda_{j+N-q_{k+j}-2} & \cdots & \Lambda_{j-q_{k+j}+q_k} & 0_{p,q_k p} \\ 0_{q_{k+j}p,p} & 0_{q_{k+j}p,p} & \cdots & 0_{q_{k+j}p,p} & 0_{q_{k+j}p,q_k p} \end{bmatrix}$$

$$q_k \coloneqq \max(k - \tilde{n}, 0)$$

Corollary 4.1 in Appendix 4.A has been used to obtain (4.11). Note that q_k is the number of $0_{p,1}$ blocks at the bottom of c_k. Observe also that $P_{\otimes,j,k}$ is a block Toeplitz matrix, as is $P_{c,j,k}$ if $q_{k+j} = q_k = 0$. Finally,

$$P_a = \mathcal{U} P_{c_\otimes} \mathcal{U}^{\mathrm{T}} = \sum_{j=0}^{\tilde{n}-1} P_{\otimes,0,j} + \sum_{i=1}^{\tilde{n}-1} \sum_{j=0}^{\tilde{n}-1-i} P^{\mathrm{F}}_{\otimes,i,j} + \left(\tilde{\mathcal{E}} \sum_{i=1}^{\tilde{n}} P_{\otimes,i,\tilde{n}-i} \right)^{\mathrm{F}} + \tilde{\mathcal{E}} P_{\otimes,0,\tilde{n}} \tilde{\mathcal{E}}^{\mathrm{T}} \quad (4.12)$$

The largest matrix that must be formed when calculating P_a using (4.12) is $P_{\otimes,j,k}$, which has dimensions $\tilde{p}^2 \times \tilde{p}^2$. Recall that (4.5) requires calculation and storage of a matrix with dimensions $(N_d \tilde{p})^2 \times (N_d \tilde{p})^2$, so (4.12) represents a significant improvement relative to (4.5). Observe however that we cannot eliminate the dependence of P_a on N_d. In (4.12), it is the summation indices which scale with N_d, rather than the memory storage requirements. With (4.12), many $P_{\otimes,j,k}$ matrices must be calculated and added together, but they need not all be stored simultaneously. Thus, as N_d increases, calculation is significantly more tractable via (4.12) than via (4.5).

Recall that P_a is not the quantity we require to calculate the optimal ALS weighting matrix. Rather, $P_{\hat{b}}$ is what we need, and by calculating $P_{\hat{b}}$ directly we can reduce the memory storage requirements further relative to substituting (4.12) into (4.4). Vectorizing (4.9) instead of (4.10) gives

$$a = (YY^{\mathrm{T}})_{\mathrm{s}} = (c_0 \otimes c_0) + (c_1 \otimes c_1) + \cdots + (c_{T_f} \otimes c_{T_f}) = \tilde{\mathcal{U}} \tilde{c}_\otimes$$

$$\tilde{\mathcal{U}} := \imath^{\mathrm{T}}_{T_f+1} \otimes I_{\tilde{p}^2} \qquad \tilde{c}_\otimes := \begin{bmatrix} c_0 \otimes c_0 \\ c_1 \otimes c_1 \\ \vdots \\ c_{T_f} \otimes c_{T_f} \end{bmatrix}$$

The covariance of $\tilde{c}_\otimes$ is

$$P_{\tilde{c}_\otimes} := \mathrm{cov}(\tilde{c}_\otimes) = \begin{bmatrix} P_{\otimes,0,0} & P^{\mathrm{T}}_{\otimes,1,0} & \cdots & P^{\mathrm{T}}_{\otimes,T_f,0} \\ P_{\otimes,1,0} & P_{\otimes,0,1} & \cdots & P^{\mathrm{T}}_{\otimes,T_f-1,1} \\ \vdots & \vdots & \ddots & \vdots \\ P_{\otimes,T_f,0} & P_{\otimes,T_f-1,1} & \cdots & P_{\otimes,0,T_f} \end{bmatrix}$$

Then P_a is the sum of all of the block elements of $P_{\tilde{c}_\otimes}$:

$$P_a = \tilde{\mathcal{U}} P_{\tilde{c}_\otimes} \tilde{\mathcal{U}}^{\mathrm{T}} = \sum_{j=0}^{T_f} P_{\otimes,0,j} + \sum_{i=1}^{T_f} \sum_{j=0}^{T_f-i} P^{\mathrm{F}}_{\otimes,i,j} \quad (4.13)$$

Next note that $\mathcal{H}$ may be expressed as

$$\mathcal{H} = \mathcal{V}\mathcal{J} \tag{4.14}$$
$$\mathcal{V} \coloneqq \mathscr{D}^{+}_{\tilde{p},p}(I_p \otimes \mathcal{F}) \qquad \mathcal{J} \coloneqq \begin{bmatrix} I_{\tilde{p}p} & 0_{\tilde{p}p,\tilde{p}p(N-1)} \end{bmatrix}$$

Substituting (4.14) into (4.4) yields

$$P_{\hat{b}} = \mathcal{V}\mathcal{J}P_a\mathcal{J}^{\mathrm{T}}\mathcal{V}^{\mathrm{T}} \tag{4.15}$$

Now we give some intermediate results that allow us to calculate $\mathcal{J}P_a\mathcal{J}^{\mathrm{T}}$.

$$\mathcal{J}(P_{c,j,k} \otimes P_{c,j,k})\mathcal{J}^{\mathrm{T}} = \Lambda_j \otimes P_{c,j,k} \tag{4.16}$$

Define $\Psi_{j,k}$ to be the first block column of $P_{c,j,k}$:

$$\Psi_{j,k} \coloneqq \begin{bmatrix} \Lambda_j \\ \Lambda_{j+1} \\ \vdots \\ \Lambda_{j+N-q_{k+j}-1} \\ 0_{q_{k+j}p,p} \end{bmatrix}$$

One can consider $\Psi_{j,k}$ a generalization of Ψ as it is defined in (3.1). In fact, $\Psi = \Psi_{0,k}$ for all $k \leq \tilde{n}$. Note that the top block row of $P_{c,j,k}$ is $\Psi^{\mathrm{T}}_{-j,k+j}$. We use this fact to obtain

$$\begin{aligned}\mathcal{J}\mathscr{K}_{\tilde{p}}(P_{c,j,k} \otimes P_{c,j,k})\mathcal{J}^{\mathrm{T}} &= \mathcal{J}\mathscr{K}_{\tilde{p}}(\Psi_{j,k} \otimes P_{c,j,k}) = \mathcal{J}(P_{c,j,k} \otimes \Psi_{j,k})\mathscr{K}_{\tilde{p},p} \\ &= (\Psi^{\mathrm{T}}_{-j,k+j} \otimes \Psi_{j,k})\mathscr{K}_{\tilde{p},p} = \mathscr{K}_{p,\tilde{p}}(\Psi_{j,k} \otimes \Psi^{\mathrm{T}}_{-j,k+j})\end{aligned} \tag{4.17}$$

Combining (4.11), (4.16), and (4.17) yields

$$\mathcal{J}P_{\otimes,,j,k}\mathcal{J}^{\mathrm{T}} = (\Lambda_j \otimes P_{c,j,k}) + \mathscr{K}_{p,\tilde{p}}(\Psi_{j,k} \otimes \Psi^{\mathrm{T}}_{-j,k+j}) \tag{4.18}$$
$$\mathcal{J}P^{\mathrm{F}}_{\otimes,,j,k}\mathcal{J}^{\mathrm{T}} = [(\Lambda_j \otimes P_{c,j,k}) + \mathscr{K}_{p,\tilde{p}}(\Psi_{j,k} \otimes \Psi^{\mathrm{T}}_{-j,k+j})]^{\mathrm{F}} \tag{4.19}$$

Finally, combining (4.13), (4.15), (4.18), and (4.19) yields

$$\begin{aligned}P_{\hat{b}} = \mathcal{V}\Big(&\sum_{j=0}^{T_f}[(\Lambda_0 \otimes P_{c,0,j}) + \mathscr{K}_{p,\tilde{p}}(\Psi_{0,j} \otimes \Psi^{\mathrm{T}}_{0,j})] \\ &+ \sum_{i=1}^{T_f}\sum_{j=0}^{T_f-i}[(\Lambda_i \otimes P_{c,i,j}) + \mathscr{K}_{p,\tilde{p}}(\Psi_{i,j} \otimes \Psi^{\mathrm{T}}_{-i,i+j})]^{\mathrm{F}}\Big)\mathcal{V}^{\mathrm{T}}\end{aligned} \tag{4.20}$$

The largest matrix that must be formed when calculating $P_{\hat{b}}$ using (4.20) has dimensions $\tilde{p}p \times \tilde{p}p$, which is a modest improvement relative to (4.12).

As N_d grows, the memory storage requirements of calculation via (4.12) and (4.20) is constant. However, calculation slows because the summation computations require many iterations. We offer two ways to speed up the calculation, both of which work by reducing the number of summation iterations required.

Recall that A is assumed to be stable, which implies that $A^k \to 0$ as $k \to \infty$. We make the approximation $A^k = 0$ for all $k \geq M$ for some M,[1] which implies

$$\begin{aligned} \Lambda_j = 0 \quad \text{and} \quad \Psi_{j,k} = 0 \quad \text{for} \quad j \geq M+1 \\ P_{c,j,k} = 0 \quad \text{and} \quad P_{\otimes,j,k} = 0 \quad \text{for} \quad j \geq M+N \end{aligned}$$

Therefore (4.12) may be rewritten as

$$P_a = \sum_{j=0}^{\tilde{n}-1} P_{\otimes,0,j} + \sum_{i=1}^{\tilde{M}_{\tilde{n}-1}} \sum_{j=0}^{\tilde{n}-1-i} P_{\otimes,i,j}^{\mathrm{F}} + \left(\tilde{\mathcal{E}} \sum_{i=1}^{\tilde{M}_{\tilde{n}}} P_{\otimes,i,\tilde{n}-i} \right)^{\mathrm{F}} + \tilde{\mathcal{E}} P_{\otimes,0,\tilde{n}} \tilde{\mathcal{E}}^{\mathrm{T}} \qquad (4.21)$$

$$\tilde{M}_k := \min(M+N-1, k)$$

Note that (4.21) is equivalent to (4.12) and (4.13) if $M \geq \tilde{n} - N + 1$. Similarly, (4.20) may be rewritten as

$$\begin{aligned} P_{\hat{b}} = \mathcal{V} \Bigg(& \sum_{j=0}^{T_f} [(\Lambda_0 \otimes P_{c,0,j}) + \mathscr{K}_{p,\tilde{p}}(\Psi_{0,j} \otimes \Psi_{0,j}^{\mathrm{T}})] \\ & + \sum_{i=1}^{M_{T_f}} \sum_{j=0}^{T_f - i} [(\Lambda_i \otimes P_{c,i,j}) + \mathscr{K}_{p,\tilde{p}}(\Psi_{i,j} \otimes \Psi_{-i,i+j}^{\mathrm{T}})]^{\mathrm{F}} \Bigg) \mathcal{V}^{\mathrm{T}} \qquad (4.22) \end{aligned}$$

$$M_k := \min(M, k)$$

and (4.22) is equivalent to (4.20) if $M \geq T_f$.

The second way to reduce the number of summation iterations is to use the fact that for $k + j \leq \tilde{n}$ and $j \geq 0$,

$$\Psi_{j,k} = \Psi_{j,0} \qquad \Psi_{-j,k+j}^{\mathrm{T}} = \Psi_{-j,0}^{\mathrm{T}} \qquad P_{c,j,k} = P_{c,j,0} \qquad P_{\otimes,j,k} = P_{\otimes,j,0}$$

[1] If A is nilpotent, i.e., $A^k = 0$ for $k \geq M_{\mathrm{nil}}$, and $M_{\mathrm{nil}} \leq M$, then this approximation is actually exact.

Therefore we may rewrite (4.21) as

$$P_a = \tilde{n}P_{\otimes,0,0} + \sum_{i=1}^{\tilde{M}_{\tilde{n}-1}} (\tilde{n}-i)P^{\mathrm{F}}_{\otimes,i,0} + \left(\tilde{\mathcal{E}}\sum_{i=1}^{\tilde{M}_{\tilde{n}}} P_{\otimes,i,0}\right)^{\mathrm{F}} + \tilde{\mathcal{E}}P_{\otimes,0,0}\tilde{\mathcal{E}}^{\mathrm{T}} \tag{4.23}$$

and (4.22) as

$$\begin{aligned} P_{\hat{b}} = \mathcal{V}\Bigg[& (\tilde{n}+1)\left[(\Lambda_0 \otimes P_{c,0,0}) + \mathscr{K}_{p,\bar{p}}(\Psi_{0,0} \otimes \Psi^{\mathrm{T}}_{0,0})\right] \\ & + \sum_{j=\tilde{n}+1}^{T_f} \left[(\Lambda_0 \otimes P_{c,0,j}) + \mathscr{K}_{p,\bar{p}}(\Psi_{0,j} \otimes \Psi^{\mathrm{T}}_{0,j})\right] \\ & + \sum_{i=1}^{M_{\tilde{n}}} \Bigg((\tilde{n}+1-i)\left[(\Lambda_i \otimes P_{c,i,0}) + \mathscr{K}_{p,\bar{p}}(\Psi_{i,0} \otimes \Psi^{\mathrm{T}}_{-i,0})\right]^{\mathrm{F}} \\ & \qquad + \sum_{j=\tilde{n}+1-i}^{T_f-i} \left[(\Lambda_i \otimes P_{c,i,j}) + \mathscr{K}_{p,\bar{p}}(\Psi_{i,j} \otimes \Psi^{\mathrm{T}}_{-i,i+j})\right]^{\mathrm{F}} \Bigg) \\ & + \sum_{i=\tilde{n}+1}^{M_{T_f}} \sum_{j=0}^{T_f-i} \left[(\Lambda_i \otimes P_{c,i,j}) + \mathscr{K}_{p,\bar{p}}(\Psi_{i,j} \otimes \Psi^{\mathrm{T}}_{-i,i+j})\right]^{\mathrm{F}} \Bigg] \mathcal{V}^{\mathrm{T}} \end{aligned} \tag{4.24}$$

Table 4.1 displays the fruits of our efforts to reduce the number of summation iterations required, giving the number of unique $P_{c,i,j}$ matrices that must be calculated in each of (4.12) and (4.20)–(4.24). For (4.21)–(4.24), the results given assume that M is small enough for any benefits to be realized. We see that in the best case scenarios, (4.23) and (4.24), the number of summation iterations required is independent of the number of data points.

When calculating P_a using (4.23), the second index in $P_{\otimes,i,j}$ is always zero, i.e., the first index specifying the lag is the only one that is consequential in the calculation. This situation arises because in (4.12), all of the $P_{\otimes,k,j}$'s that are needed do not have any zero blocks in them, as the indices always correspond to $q_{k+j} = q_k = 0$. This result is a consequence of using the $\mathcal{E}_k$ matrices to account for c_k for $k \geq \tilde{n}$, the indices for which zero blocks are present. When calculating $P_{\hat{b}}$ directly, as in (4.24), it is not possible to account for c_k for $k \geq \tilde{n}$ in this manner, so the second index is necessary. Equation (4.23) is arguably more convenient to use than (4.24), since the former has only one index which must be tracked. The advantage of (4.24) is moderately smaller memory storage requirements.

Equation	Assumption on M	Unique $P_{c,i,j}$'s
(4.12)	none	$\frac{(\tilde{n}+1)(\tilde{n}+2)}{2}$
(4.20)	none	$\frac{(\tilde{n}+N)(\tilde{n}+N+1)}{2}$
(4.21)	$M \leq \tilde{n} - N$	$\frac{(M+N)(2\tilde{n}+3-M-N)}{2}$
(4.22)	$M \leq \tilde{n} + N - 1$	$\frac{(M+1)(2\tilde{n}+2N-M)}{2}$
(4.23)	$M \leq \tilde{n} - N$	$M + N$
(4.24)	$M \leq \tilde{n}$	$N(M+1)$

Table 4.1: Number of unique $P_{c,i,j}$ matrices that are required for calculation of P_a and $P_{\hat{b}}$ using various equations.

Appendices

4.A Miscellaneous results

Lemma 4.1. *Let n and m be positive integers. Then*

$$\begin{bmatrix} \mathcal{K}_{n+m,n} & 0 \\ 0 & \mathcal{K}_{n+m,m} \end{bmatrix} \mathcal{K}_{n+m} \begin{bmatrix} \mathcal{K}_{n,n+m} & 0 \\ 0 & \mathcal{K}_{m,n+m} \end{bmatrix} = \begin{bmatrix} \mathcal{K}_n & 0 & 0 & 0 \\ 0 & 0 & \mathcal{K}_{m,n} & 0 \\ 0 & \mathcal{K}_{n,m} & 0 & 0 \\ 0 & 0 & 0 & \mathcal{K}_m \end{bmatrix}$$

Proof. Consider an arbitrary block matrix

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

with $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{m\times n}$, and $D \in \mathbb{R}^{m\times m}$. Observe the following

identity:

$$
\begin{aligned}
&\begin{bmatrix} \mathscr{K}_{n+m,n} & 0 \\ 0 & \mathscr{K}_{n+m,m} \end{bmatrix} \mathscr{K}_{n+m} \begin{bmatrix} \mathscr{K}_{n,n+m} & 0 \\ 0 & \mathscr{K}_{m,n+m} \end{bmatrix} \begin{bmatrix} A_{\mathrm{s}} \\ B_{\mathrm{s}} \\ C_{\mathrm{s}} \\ D_{\mathrm{s}} \end{bmatrix} \\
&\quad = \begin{bmatrix} \mathscr{K}_{n+m,n} & 0 \\ 0 & \mathscr{K}_{n+m,m} \end{bmatrix} \mathscr{K}_{n+m} \begin{bmatrix} \mathscr{K}_{n,n+m} & 0 \\ 0 & \mathscr{K}_{m,n+m} \end{bmatrix} \begin{bmatrix} \begin{bmatrix} A & B \end{bmatrix}_{\mathrm{s}} \\ \begin{bmatrix} C & D \end{bmatrix}_{\mathrm{s}} \end{bmatrix} \\
&\quad = \begin{bmatrix} \mathscr{K}_{n+m,n} & 0 \\ 0 & \mathscr{K}_{n+m,m} \end{bmatrix} \mathscr{K}_{n+m} \begin{bmatrix} \mathscr{K}_{n,n+m} \begin{bmatrix} A & B \end{bmatrix}_{\mathrm{s}} \\ \mathscr{K}_{m,n+m} \begin{bmatrix} C & D \end{bmatrix}_{\mathrm{s}} \end{bmatrix} \\
&\quad = \begin{bmatrix} \mathscr{K}_{n+m,n} & 0 \\ 0 & \mathscr{K}_{n+m,m} \end{bmatrix} \mathscr{K}_{n+m} \begin{bmatrix} \left(\begin{bmatrix} A & B \end{bmatrix}^{\mathrm{T}} \right)_{\mathrm{s}} \\ \left(\begin{bmatrix} C & D \end{bmatrix}^{\mathrm{T}} \right)_{\mathrm{s}} \end{bmatrix} \\
&\quad = \begin{bmatrix} \mathscr{K}_{n+m,n} & 0 \\ 0 & \mathscr{K}_{n+m,m} \end{bmatrix} \mathscr{K}_{n+m} \begin{bmatrix} \begin{bmatrix} A^{\mathrm{T}} \\ B^{\mathrm{T}} \end{bmatrix}_{\mathrm{s}} \\ \begin{bmatrix} C^{\mathrm{T}} \\ D^{\mathrm{T}} \end{bmatrix}_{\mathrm{s}} \end{bmatrix} \\
&\quad = \begin{bmatrix} \mathscr{K}_{n+m,n} & 0 \\ 0 & \mathscr{K}_{n+m,m} \end{bmatrix} \mathscr{K}_{n+m} \begin{bmatrix} A^{\mathrm{T}} & C^{\mathrm{T}} \\ B^{\mathrm{T}} & D^{\mathrm{T}} \end{bmatrix}_{\mathrm{s}} \\
&\quad = \begin{bmatrix} \mathscr{K}_{n+m,n} & 0 \\ 0 & \mathscr{K}_{n+m,m} \end{bmatrix} \mathscr{K}_{n+m} \left(\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{\mathrm{T}} \right)_{\mathrm{s}} \\
&\quad = \begin{bmatrix} \mathscr{K}_{n+m,n} & 0 \\ 0 & \mathscr{K}_{n+m,m} \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix}_{\mathrm{s}} \\
&\quad = \begin{bmatrix} \mathscr{K}_{n+m,n} & 0 \\ 0 & \mathscr{K}_{n+m,m} \end{bmatrix} \begin{bmatrix} \begin{bmatrix} A \\ C \end{bmatrix}_{\mathrm{s}} \\ \begin{bmatrix} B \\ D \end{bmatrix}_{\mathrm{s}} \end{bmatrix} \\
&\quad = \begin{bmatrix} \mathscr{K}_{n+m,n} \begin{bmatrix} A \\ C \end{bmatrix}_{\mathrm{s}} \\ \mathscr{K}_{n+m,m} \begin{bmatrix} B \\ D \end{bmatrix}_{\mathrm{s}} \end{bmatrix} \\
&\quad = \begin{bmatrix} \mathscr{K}_{n+m,n} \left(\begin{bmatrix} A^{\mathrm{T}} & C^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} \right)_{\mathrm{s}} \\ \mathscr{K}_{n+m,m} \left(\begin{bmatrix} B^{\mathrm{T}} & D^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} \right)_{\mathrm{s}} \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
&= \begin{bmatrix} \begin{bmatrix} A^{\mathrm{T}} & C^{\mathrm{T}} \end{bmatrix}_{\mathrm{s}} \\ \begin{bmatrix} B^{\mathrm{T}} & D^{\mathrm{T}} \end{bmatrix}_{\mathrm{s}} \end{bmatrix} \\
&= \begin{bmatrix} (A^{\mathrm{T}})_{\mathrm{s}} \\ (C^{\mathrm{T}})_{\mathrm{s}} \\ (B^{\mathrm{T}})_{\mathrm{s}} \\ (D^{\mathrm{T}})_{\mathrm{s}} \end{bmatrix} \\
&= \begin{bmatrix} \mathscr{K}_n A_{\mathrm{s}} \\ \mathscr{K}_{m,n} C_{\mathrm{s}} \\ \mathscr{K}_{n,m} B_{\mathrm{s}} \\ \mathscr{K}_m D_{\mathrm{s}} \end{bmatrix} \\
&= \begin{bmatrix} \mathscr{K}_n & 0 & 0 & 0 \\ 0 & 0 & \mathscr{K}_{m,n} & 0 \\ 0 & \mathscr{K}_{n,m} & 0 & 0 \\ 0 & 0 & 0 & \mathscr{K}_m \end{bmatrix} \begin{bmatrix} A_{\mathrm{s}} \\ B_{\mathrm{s}} \\ C_{\mathrm{s}} \\ D_{\mathrm{s}} \end{bmatrix}
\end{aligned}
$$

The result follows because A, B, C, and D are arbitrary (aside from the fact that they must have conforming dimensions). ■

Theorem 4.1. *Let $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$ be jointly distributed Gaussian random variables:*

$$
\begin{bmatrix} x \\ y \end{bmatrix} \sim \mathrm{N}\left(\begin{bmatrix} m_x \\ m_y \end{bmatrix}, \begin{bmatrix} P_x & P_{xy} \\ P_{yx} & P_y \end{bmatrix} \right)
$$

Then

$$
\mathrm{E} \begin{bmatrix} x \otimes x \\ y \otimes x \\ x \otimes y \\ y \otimes y \end{bmatrix} = \begin{bmatrix} (P_x)_{\mathrm{s}} \\ (P_{xy})_{\mathrm{s}} \\ (P_{yx})_{\mathrm{s}} \\ (P_y)_{\mathrm{s}} \end{bmatrix} + \begin{bmatrix} m_x \otimes m_x \\ m_y \otimes m_x \\ m_x \otimes m_y \\ m_y \otimes m_y \end{bmatrix} \qquad \operatorname{cov} \begin{bmatrix} x \otimes x \\ y \otimes x \\ x \otimes y \\ y \otimes y \end{bmatrix} = K(P_1 + P_2 + P_3)
$$

where

$$
K \coloneqq \begin{bmatrix} I_{n^2} + \mathscr{K}_n & 0 & 0 & 0 \\ 0 & I_{nm} & \mathscr{K}_{m,n} & 0 \\ 0 & \mathscr{K}_{n,m} & I_{nm} & 0 \\ 0 & 0 & 0 & I_{m^2} + \mathscr{K}_m \end{bmatrix}
$$

$$
P_1 \coloneqq \begin{bmatrix} P_x \otimes P_x & P_{xy} \otimes P_x & P_x \otimes P_{xy} & P_{xy} \otimes P_{xy} \\ P_{yx} \otimes P_x & P_y \otimes P_x & P_{yx} \otimes P_{xy} & P_y \otimes P_{xy} \\ P_x \otimes P_{yx} & P_{xy} \otimes P_{yx} & P_x \otimes P_y & P_{xy} \otimes P_y \\ P_{yx} \otimes P_{yx} & P_y \otimes P_{yx} & P_{yx} \otimes P_y & P_y \otimes P_y \end{bmatrix}
$$

$$P_2 := \begin{bmatrix} m_x m_x^{\mathrm{T}} \otimes P_x & m_x m_y^{\mathrm{T}} \otimes P_x & m_x m_x^{\mathrm{T}} \otimes P_{xy} & m_x m_y^{\mathrm{T}} \otimes P_{xy} \\ m_y m_x^{\mathrm{T}} \otimes P_x & m_y m_y^{\mathrm{T}} \otimes P_x & m_y m_x^{\mathrm{T}} \otimes P_{xy} & m_y m_y^{\mathrm{T}} \otimes P_{xy} \\ m_x m_x^{\mathrm{T}} \otimes P_{yx} & m_x m_y^{\mathrm{T}} \otimes P_{yx} & m_x m_x^{\mathrm{T}} \otimes P_y & m_x m_y^{\mathrm{T}} \otimes P_y \\ m_y m_x^{\mathrm{T}} \otimes P_{yx} & m_y m_y^{\mathrm{T}} \otimes P_{yx} & m_y m_x^{\mathrm{T}} \otimes P_y & m_y m_y^{\mathrm{T}} \otimes P_y \end{bmatrix}$$

$$P_3 := \begin{bmatrix} P_x \otimes m_x m_x^{\mathrm{T}} & P_{xy} \otimes m_x m_x^{\mathrm{T}} & P_x \otimes m_x m_y^{\mathrm{T}} & P_{xy} \otimes m_x m_y^{\mathrm{T}} \\ P_{yx} \otimes m_x m_x^{\mathrm{T}} & P_y \otimes m_x m_x^{\mathrm{T}} & P_{yx} \otimes m_x m_y^{\mathrm{T}} & P_y \otimes m_x m_y^{\mathrm{T}} \\ P_x \otimes m_y m_x^{\mathrm{T}} & P_{xy} \otimes m_y m_x^{\mathrm{T}} & P_x \otimes m_y m_y^{\mathrm{T}} & P_{xy} \otimes m_y m_y^{\mathrm{T}} \\ P_{yx} \otimes m_y m_x^{\mathrm{T}} & P_y \otimes m_y m_x^{\mathrm{T}} & P_{yx} \otimes m_y m_y^{\mathrm{T}} & P_y \otimes m_y m_y^{\mathrm{T}} \end{bmatrix}$$

Proof. We need some preliminary results before we proceed with the proof. First, note the following linear relationship:

$$\begin{bmatrix} x \\ y \end{bmatrix} \otimes \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \otimes \begin{bmatrix} x \\ y \end{bmatrix} \\ y \otimes \begin{bmatrix} x \\ y \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \mathscr{K}_{n,n+m}\left(\begin{bmatrix} x \\ y \end{bmatrix} \otimes x\right) \\ \mathscr{K}_{m,n+m}\left(\begin{bmatrix} x \\ y \end{bmatrix} \otimes y\right) \end{bmatrix} = \begin{bmatrix} \mathscr{K}_{n,n+m} & 0 \\ 0 & \mathscr{K}_{m,n+m} \end{bmatrix} \begin{bmatrix} x \otimes x \\ y \otimes x \\ x \otimes y \\ y \otimes y \end{bmatrix}$$

The inverse transformation is

$$\begin{bmatrix} x \otimes x \\ y \otimes x \\ x \otimes y \\ y \otimes y \end{bmatrix} = \begin{bmatrix} \mathscr{K}_{n+m,n} & 0 \\ 0 & \mathscr{K}_{n+m,m} \end{bmatrix} \left(\begin{bmatrix} x \\ y \end{bmatrix} \otimes \begin{bmatrix} x \\ y \end{bmatrix} \right) \tag{4.25}$$

Next we establish three identities. First,

$$\begin{aligned} \begin{bmatrix} P_x & P_{xy} \\ P_{yx} & P_y \end{bmatrix} \otimes \begin{bmatrix} P_x & P_{xy} \\ P_{yx} & P_y \end{bmatrix} &= \begin{bmatrix} P_x \otimes \begin{bmatrix} P_x & P_{xy} \\ P_{yx} & P_y \end{bmatrix} & P_{xy} \otimes \begin{bmatrix} P_x & P_{xy} \\ P_{yx} & P_y \end{bmatrix} \\ P_{yx} \otimes \begin{bmatrix} P_x & P_{xy} \\ P_{yx} & P_y \end{bmatrix} & P_y \otimes \begin{bmatrix} P_x & P_{xy} \\ P_{yx} & P_y \end{bmatrix} \end{bmatrix} \\ &= \begin{bmatrix} \mathscr{K}_{n,n+m}\left(\begin{bmatrix} P_x & P_{xy} \\ P_{yx} & P_y \end{bmatrix} \otimes P_x\right)\mathscr{K}_{n+m,n} & \mathscr{K}_{n,n+m}\left(\begin{bmatrix} P_x & P_{xy} \\ P_{yx} & P_y \end{bmatrix} \otimes P_{xy}\right)\mathscr{K}_{n+m,m} \\ \mathscr{K}_{m,n+m}\left(\begin{bmatrix} P_x & P_{xy} \\ P_{yx} & P_y \end{bmatrix} \otimes P_{yx}\right)\mathscr{K}_{n+m,n} & \mathscr{K}_{m,n+m}\left(\begin{bmatrix} P_x & P_{xy} \\ P_{yx} & P_y \end{bmatrix} \otimes P_y\right)\mathscr{K}_{n+m,m} \end{bmatrix} \\ &= \begin{bmatrix} \mathscr{K}_{n,n+m} & 0 \\ 0 & \mathscr{K}_{m,n+m} \end{bmatrix} P_1 \begin{bmatrix} \mathscr{K}_{n+m,n} & 0 \\ 0 & \mathscr{K}_{n+m,m} \end{bmatrix} \end{aligned} \tag{4.26}$$

Similarly,

$$\begin{bmatrix} P_x & P_{xy} \\ P_{yx} & P_y \end{bmatrix} \otimes \begin{bmatrix} m_x m_x^{\mathrm{T}} & m_x m_y^{\mathrm{T}} \\ m_y m_x^{\mathrm{T}} & m_y m_y^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} \mathscr{K}_{n,n+m} & 0 \\ 0 & \mathscr{K}_{m,n+m} \end{bmatrix} P_2 \begin{bmatrix} \mathscr{K}_{n+m,n} & 0 \\ 0 & \mathscr{K}_{n+m,m} \end{bmatrix} \tag{4.27}$$

$$\begin{bmatrix} m_x m_x^{\mathrm{T}} & m_x m_y^{\mathrm{T}} \\ m_y m_x^{\mathrm{T}} & m_y m_y^{\mathrm{T}} \end{bmatrix} \otimes \begin{bmatrix} P_x & P_{xy} \\ P_{yx} & P_y \end{bmatrix} = \begin{bmatrix} \mathscr{K}_{n,n+m} & 0 \\ 0 & \mathscr{K}_{m,n+m} \end{bmatrix} P_3 \begin{bmatrix} \mathscr{K}_{n+m,n} & 0 \\ 0 & \mathscr{K}_{n+m,m} \end{bmatrix} \tag{4.28}$$

Using results from Ghazal and Neudecker (2000, Section 3), we have

$$\begin{aligned} \mathrm{E}\left(\begin{bmatrix} x \\ y \end{bmatrix} \otimes \begin{bmatrix} x \\ y \end{bmatrix}\right) &= \begin{bmatrix} P_x & P_{xy} \\ P_{yx} & P_y \end{bmatrix}_{\mathrm{s}} + \begin{bmatrix} m_x \\ m_y \end{bmatrix} \otimes \begin{bmatrix} m_x \\ m_y \end{bmatrix} \\ &= \begin{bmatrix} P_x & P_{xy} \\ P_{yx} & P_y \end{bmatrix}_{\mathrm{s}} + \begin{bmatrix} m_x \otimes \begin{bmatrix} m_x \\ m_y \end{bmatrix} \\ m_y \otimes \begin{bmatrix} m_x \\ m_y \end{bmatrix} \end{bmatrix} = \begin{bmatrix} P_x & P_{xy} \\ P_{yx} & P_y \end{bmatrix}_{\mathrm{s}} + \begin{bmatrix} \mathscr{K}_{n,n+m}\left(\begin{bmatrix} m_x \\ m_y \end{bmatrix} \otimes m_x\right) \\ \mathscr{K}_{m,n+m}\left(\begin{bmatrix} m_x \\ m_y \end{bmatrix} \otimes m_y\right) \end{bmatrix} \\ &= \begin{bmatrix} P_x & P_{xy} \\ P_{yx} & P_y \end{bmatrix}_{\mathrm{s}} + \begin{bmatrix} \mathscr{K}_{n,n+m} & 0 \\ 0 & \mathscr{K}_{m,n+m} \end{bmatrix} \begin{bmatrix} m_x \otimes m_x \\ m_y \otimes m_x \\ m_x \otimes m_y \\ m_y \otimes m_y \end{bmatrix} \end{aligned} \tag{4.29}$$

and

$$\begin{aligned} \operatorname{cov}\left(\begin{bmatrix} x \\ y \end{bmatrix} \otimes \begin{bmatrix} x \\ y \end{bmatrix}\right) &= (I_{(n+m)^2} + \mathscr{K}_{n+m})\left(\begin{bmatrix} P_x & P_{xy} \\ P_{yx} & P_y \end{bmatrix} \otimes \begin{bmatrix} P_x & P_{xy} \\ P_{yx} & P_y \end{bmatrix}\right. \\ &\qquad \left. + \begin{bmatrix} P_x & P_{xy} \\ P_{yx} & P_y \end{bmatrix} \otimes \begin{bmatrix} m_x m_x^{\mathrm{T}} & m_x m_y^{\mathrm{T}} \\ m_y m_x^{\mathrm{T}} & m_y m_y^{\mathrm{T}} \end{bmatrix} + \begin{bmatrix} m_x m_x^{\mathrm{T}} & m_x m_y^{\mathrm{T}} \\ m_y m_x^{\mathrm{T}} & m_y m_y^{\mathrm{T}} \end{bmatrix} \otimes \begin{bmatrix} P_x & P_{xy} \\ P_{yx} & P_y \end{bmatrix}\right) \\ &= (I_{(n+m)^2} + \mathscr{K}_{n+m}) \begin{bmatrix} \mathscr{K}_{n,n+m} & 0 \\ 0 & \mathscr{K}_{m,n+m} \end{bmatrix} (P_1 + P_2 + P_3) \begin{bmatrix} \mathscr{K}_{n+m,n} & 0 \\ 0 & \mathscr{K}_{n+m,m} \end{bmatrix} \end{aligned} \tag{4.30}$$

Equations (4.26)–(4.28) have been used to arrive at (4.30). Now we derive the expectation. Using (4.25), (4.29), and the linearity of expectation, we obtain

$$\begin{aligned} \mathrm{E} \begin{bmatrix} x \otimes x \\ y \otimes x \\ x \otimes y \\ y \otimes y \end{bmatrix} &= \begin{bmatrix} \mathscr{K}_{n+m,n} & 0 \\ 0 & \mathscr{K}_{n+m,m} \end{bmatrix} \mathrm{E}\left(\begin{bmatrix} x \\ y \end{bmatrix} \otimes \begin{bmatrix} x \\ y \end{bmatrix}\right) \\ &= \begin{bmatrix} \mathscr{K}_{n+m,n} & 0 \\ 0 & \mathscr{K}_{n+m,m} \end{bmatrix} \left(\begin{bmatrix} P_x & P_{xy} \\ P_{yx} & P_y \end{bmatrix}_{\mathrm{s}} + \begin{bmatrix} \mathscr{K}_{n,n+m} & 0 \\ 0 & \mathscr{K}_{m,n+m} \end{bmatrix} \begin{bmatrix} m_x \otimes m_x \\ m_y \otimes m_x \\ m_x \otimes m_y \\ m_y \otimes m_y \end{bmatrix} \right) \end{aligned}$$

$$= \begin{bmatrix} \mathcal{K}_{n+m,n} \begin{bmatrix} P_x \\ P_{yx} \end{bmatrix}_{\mathrm{s}} \\ \mathcal{K}_{n+m,m} \begin{bmatrix} P_{xy} \\ P_y \end{bmatrix}_{\mathrm{s}} \end{bmatrix} + \begin{bmatrix} m_x \otimes m_x \\ m_y \otimes m_x \\ m_x \otimes m_y \\ m_y \otimes m_y \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} P_x & P_{xy} \end{bmatrix}_{\mathrm{s}} \\ \begin{bmatrix} P_{yx} & P_y \end{bmatrix}_{\mathrm{s}} \end{bmatrix} + \begin{bmatrix} m_x \otimes m_x \\ m_y \otimes m_x \\ m_x \otimes m_y \\ m_y \otimes m_y \end{bmatrix}$$
$$= \begin{bmatrix} (P_x)_{\mathrm{s}} \\ (P_{xy})_{\mathrm{s}} \\ (P_{yx})_{\mathrm{s}} \\ (P_y)_{\mathrm{s}} \end{bmatrix} + \begin{bmatrix} m_x \otimes m_x \\ m_y \otimes m_x \\ m_x \otimes m_y \\ m_y \otimes m_y \end{bmatrix}$$

Now we derive the covariance. Using (4.25), (4.30), and Lemma 4.1, we obtain

$$\operatorname{cov} \begin{bmatrix} x \otimes x \\ y \otimes x \\ x \otimes y \\ y \otimes y \end{bmatrix} = \begin{bmatrix} \mathcal{K}_{n+m,n} & 0 \\ 0 & \mathcal{K}_{n+m,m} \end{bmatrix} \operatorname{cov}\left(\begin{bmatrix} x \\ y \end{bmatrix} \otimes \begin{bmatrix} x \\ y \end{bmatrix} \right) \begin{bmatrix} \mathcal{K}_{n,n+m} & 0 \\ 0 & \mathcal{K}_{m,n+m} \end{bmatrix}$$
$$= \begin{bmatrix} \mathcal{K}_{n+m,n} & 0 \\ 0 & \mathcal{K}_{n+m,m} \end{bmatrix} (I_{(n+m)^2} + \mathcal{K}_{n+m}) \begin{bmatrix} \mathcal{K}_{n,n+m} & 0 \\ 0 & \mathcal{K}_{m,n+m} \end{bmatrix} (P_1 + P_2 + P_3)$$
$$\cdot \begin{bmatrix} \mathcal{K}_{n+m,n} & 0 \\ 0 & \mathcal{K}_{n+m,m} \end{bmatrix} \begin{bmatrix} \mathcal{K}_{n,n+m} & 0 \\ 0 & \mathcal{K}_{m,n+m} \end{bmatrix}$$
$$= \left(I_{(n+m)^2} + \begin{bmatrix} \mathcal{K}_{n+m,n} & 0 \\ 0 & \mathcal{K}_{n+m,m} \end{bmatrix} \mathcal{K}_{n+m} \begin{bmatrix} \mathcal{K}_{n,n+m} & 0 \\ 0 & \mathcal{K}_{m,n+m} \end{bmatrix} \right) (P_1 + P_2 + P_3)$$
$$= \left(I_{(n+m)^2} + \begin{bmatrix} \mathcal{K}_n & 0 & 0 & 0 \\ 0 & 0 & \mathcal{K}_{m,n} & 0 \\ 0 & \mathcal{K}_{n,m} & 0 & 0 \\ 0 & 0 & 0 & \mathcal{K}_m \end{bmatrix} \right) (P_1 + P_2 + P_3)$$
$$= K(P_1 + P_2 + P_3) \qquad \blacksquare$$

Corollary 4.1. *Let $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$ be jointly distributed Gaussian random variables:*

$$\begin{bmatrix} x \\ y \end{bmatrix} \sim \mathrm{N}\left(\begin{bmatrix} m_x \\ m_y \end{bmatrix}, \begin{bmatrix} P_x & P_{xy} \\ P_{yx} & P_y \end{bmatrix} \right)$$

Then

$$\operatorname{cov}(x \otimes x, y \otimes y) = (I_{n^2} + \mathcal{K}_n)(P_{xy} \otimes P_{xy} + m_x m_y^{\mathrm{T}} \otimes P_{xy} + P_{xy} \otimes m_x m_y^{\mathrm{T}})$$
$$\operatorname{cov}(y \otimes y, x \otimes x) = (I_{m^2} + \mathcal{K}_m)(P_{yx} \otimes P_{yx} + m_y m_x^{\mathrm{T}} \otimes P_{yx} + P_{yx} \otimes m_y m_x^{\mathrm{T}})$$

Proof. These matrices are the upper right and lower left blocks of

$$\operatorname{cov}\begin{bmatrix} x \otimes x \\ y \otimes x \\ x \otimes y \\ y \otimes y \end{bmatrix}$$

respectively. ■

Even at your best, the creative moments are still kind of fleeting.

Rivers Cuomo

CHAPTER 5

MAXIMUM LIKELIHOOD ESTIMATION

This chapter introduces the maximum likelihood (ML) technique and applies it for covariance estimation. ML is a general method for statistical parameter estimation in which the likelihood function of the data is maximized. That is, if a data set y has probability distribution $p(y;\theta)$, where θ is a parameter, the ML estimate of θ is

$$\hat{\theta}_{\mathrm{ML}} = \arg\max_{\theta} p(y;\theta) \tag{5.1}$$

We present two slightly different versions of the ML problem to estimate the covariance matrices Q and R under the assumption $S = 0$. As with Chapter 4, here we consider only the process system without inputs and A stable, (1.3) and (1.4), and not the $\mathcal{K}$-error system, (2.10) and (2.11). Additionally, unlike for ALS, ML-based methods require that the distribution of the noises be known (or assumed). Therefore, we again assume that (w_k) and (v_k) are jointly normally distributed, as in (4.1). One way in which we break from previous chapters is that we study ML estimation for $(Q, R, 0)$, but not for (Q, R, S), (Φ, Θ), or $(\mathbb{\Phi}, \mathbb{\Theta})$.

The chapter is organized as follows:

- Section 5.1 introduces the ML technique for covariance estimation, presenting a version of the problem that we refer to as direct ML estimation.
- Section 5.2 discusses the expectation maximization (EM) algorithm for solving ML problems, and presents an ML problem for covariance estimation to which EM is well suited for solving.

- Section 5.3 discusses the literature of ML methods for covariance estimation.

The chapter also contains one appendix: 5.A.

5.1 Direct maximum likelihood estimation

First, we derive the likelihood function for Q and R. Assume that the outputs sequence $y_{0:T_f}$ is Gaussian and stationary, i.e., (2.21)–(2.24) hold. Recall that $\tilde{y}$ is the vectorized version of $y_{0:T_f}$ (see (4.6)). The vector $\tilde{y}$ is a linear function of x_0, $w_{0:T_f-1}$ and $v_{0:T_f}$:

$$\tilde{y} = \mathcal{O}_{1,T_f+1}x_0 + \mathcal{O}_{3,T_f}(I_{T_f} \otimes G)\tilde{w} + (I_{T_f+1} \otimes H)\tilde{v}$$

$$\mathcal{O}_{3,N} := \begin{bmatrix} 0_{p,n} & & & & \\ C & 0_{p,n} & & & \\ CA & C & & & \\ \vdots & \vdots & \ddots & & \\ CA^{N-3} & CA^{N-4} & \cdots & 0_{p,n} & \\ CA^{N-2} & CA^{N-3} & \cdots & C & 0_{p,n} \\ CA^{N-1} & CA^{N-2} & \cdots & CA & C \end{bmatrix} \qquad \tilde{w} := \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_{T_f-1} \end{bmatrix} \qquad \tilde{v} := \begin{bmatrix} v_0 \\ v_1 \\ \vdots \\ v_{T_f} \end{bmatrix}$$

From (2.20), $x_0 \sim \mathrm{N}(0, P)$ where P is the unique solution to (2.18), $P = APA^{\mathrm{T}} + GQG^{\mathrm{T}}$. Hence

$$\tilde{y} \sim \mathrm{N}(0, P_{\tilde{y}})$$
$$P_{\tilde{y}} := \mathcal{O}_{1,T_f+1}P\mathcal{O}_{1,T_f+1}^{\mathrm{T}} + \mathcal{O}_{3,T_f}(I_{T_f} \otimes GQG^{\mathrm{T}})\mathcal{O}_{3,T_f}^{\mathrm{T}} + (I_{T_f+1} \otimes HRH^{\mathrm{T}})$$

Note that $P_{\tilde{y}}$ is a function of (Q, R), and therefore it might be more properly written as $P_{\tilde{y}}(Q, R)$. However, for the sake of convenience and readability, we do not denote explicitly this functional dependence. The probability density function of $\tilde{y}$ is the normal distribution:

$$p_{\tilde{y}}(\tilde{y}; Q, R) = \frac{\exp(-\frac{1}{2}\tilde{y}^{\mathrm{T}}P_{\tilde{y}}^{-1}\tilde{y})}{(2\pi)^{(T_f+1)p/2}|P_{\tilde{y}}|^{1/2}} \tag{5.2}$$

The log-likelihood function is

$$\ln p_{\tilde{y}}(\tilde{y}; Q, R) = -\frac{(T_f+1)p}{2}\ln 2\pi - \frac{1}{2}\ln|P_{\tilde{y}}| - \frac{1}{2}\tilde{y}^{\mathrm{T}}P_{\tilde{y}}^{-1}\tilde{y}$$

The ML problem is

$$\begin{aligned}(\hat{Q},\hat{R}) &= \arg\max_{Q,R} \ \ln p_{\tilde{y}}(\tilde{y};Q,R) \quad \text{s.t. } Q \geq 0,\ R \geq 0 \\ &= \arg\min_{Q,R} \ \phi(Q,R) \quad \text{s.t. } Q \geq 0,\ R \geq 0 \\ \phi(Q,R) &\coloneqq -2\ln p_{\tilde{y}}(\tilde{y};Q,R) + (T_f+1)p\ln 2\pi \\ &= \ln|P_{\tilde{y}}| + \tilde{y}^{\mathrm{T}}P_{\tilde{y}}^{-1}\tilde{y}\end{aligned} \tag{5.3}$$

We refer to (5.3) as the direct ML problem. This problem implicitly assumes that the inverse $P_{\tilde{y}}^{-1}$ exists, since it appears in the objective function ϕ.

The matrix $P_{\tilde{y}}$ has dimensions $(T_f+1)p \times (T_f+1)p$. Evaluating the ML objective function $\phi(Q,R)$ requires calculating the log determinant and the inverse of this matrix. The computational expense of these operations scales rapidly with the size of $P_{\tilde{y}}$. As a result, direct solution of (5.3) using a general nonlinear programming solver is usually intractable for all but small to moderately sized problems. Two strategies which can be used to improve tractability are to assume Q and R are diagonal so there are fewer optimization variables and to assume $A^k = 0$ for $k \geq M$ for some M and form $P_{\tilde{y}}$ as a sparse matrix. We discuss these strategies further and evaluate their efficacy in speeding ML computation in Chapter 7.

5.2 Maximum likelihood via expectation maximization

Expectation maximization (EM) is a general algorithm for solving ML problems that was first formalized in a general manner by Dempster et al. (1977), although others had earlier developed versions of the algorithm for specific applications. Wu (1983) offered some important corrections regarding convergence properties of EM. A high level explanation of EM is given by Do and Batzoglou (2008), and Borman (2004) provides a brief and accessible tutorial explaining some of the details of the technique. McLachlan and Krishnan (2008) give a comprehensive treatment of all aspects of the EM algorithm.

For the general ML problem (5.1), EM relies on the notion that there are hidden variables, denoted x, in addition to the observed variables y, and that the joint density of x and y is parameterized by θ, i.e., $p(y,x;\theta)$. With covariance estimation, the observed variables are the outputs, $y_{0:T_f}$, and the hidden variables are the states, $x_{0:T_f}$. An EM iteration for (5.1) is given by

$$\hat{\theta}_{k+1} = \arg\max_{\theta} \mathrm{E}_{x|y;\hat{\theta}_k}\left[\ln p(y,x;\theta)\right] \tag{5.4}$$

The notation $\mathrm{E}_{x|y;\hat{\theta}_k}[\cdot]$ means to take the conditional expectation under the assumption that $\hat{\theta}_k$ parameterizes the conditional probability distribution. Authors often discuss the expectation and maximization steps separately, although we have combined them in (5.4).

We summarize the most important properties of EM below:

1. The likelihood function is nondecreasing for every EM iteration (5.4), i.e., $p(y;\theta_{k+1}) \geq p(y;\theta_k)$.

2. Under general conditions, EM converges to a stationary point of the likelihood function. In practice, this stationary point is nearly always a local maximum of $p(y;\theta)$, but it is possible that it can be a saddle point or a local minimum.

3. The convergence rate of EM is usually slow, but the individual iterations can often be performed rapidly.

4. EM does not anywhere require evaluation of the likelihood function $p(y;\theta)$ (recall that this evaluation is the difficulty with direct solution of (5.3)).

Due to the difficulty of directly solving (5.3), most work on the covariance estimation ML problem has focused on using the EM algorithm as a means of solution. However, EM is amenable to a slightly different form of the ML problem than (5.3). Instead of assuming the outputs sequence has the stationary distribution (2.21)–(2.24), assume $x_0 \sim \mathrm{N}(\mu, P_0)$. Then

$$
\begin{aligned}
\tilde{y} &\sim \mathrm{N}(\mathcal{O}_{1,T_f+1}\mu, \tilde{P}_{\tilde{y}}) \\
\tilde{P}_{\tilde{y}} &\coloneqq \mathcal{O}_{1,T_f+1} P_0 \mathcal{O}_{1,T_f+1}^{\mathrm{T}} + \mathcal{O}_{3,T_f}(I_{T_f} \otimes GQG^{\mathrm{T}})\mathcal{O}_{3,T_f}^{\mathrm{T}} + (I_{T_f+1} \otimes HRH^{\mathrm{T}}) \\
p_{\tilde{y}}(\tilde{y};\mu,P_0,Q,R) &= \frac{\exp[-\frac{1}{2}(\tilde{y}-\mathcal{O}_{1,T_f+1}\mu)^{\mathrm{T}}\tilde{P}_{\tilde{y}}^{-1}(\tilde{y}-\mathcal{O}_{1,T_f+1}\mu)]}{(2\pi)^{(T_f+1)p/2}|\tilde{P}_{\tilde{y}}|^{1/2}} \\
\ln p_{\tilde{y}}(\tilde{y};\mu,P_0,Q,R) &= -\tfrac{(T_f+1)p}{2}\ln 2\pi - \tfrac{1}{2}\ln|\tilde{P}_{\tilde{y}}| \\
&\quad -\tfrac{1}{2}(\tilde{y}-\mathcal{O}_{1,T_f+1}\mu)^{\mathrm{T}}\tilde{P}_{\tilde{y}}^{-1}(\tilde{y}-\mathcal{O}_{1,T_f+1}\mu)
\end{aligned}
$$

The ML problem becomes

$$\begin{aligned}(\hat{\mu}, \hat{P}_0, \hat{Q}, \hat{R}) &= \arg \max_{\mu, P_0, Q, R} \ln p_{\tilde{y}}(\tilde{y}; \mu, P_0, Q, R) \quad \text{s.t. } P_0 \geq 0,\ Q \geq 0,\ R \geq 0 \\ &= \arg \min_{\mu, P_0, Q, R} \tilde{\phi}(\mu, P_0, Q, R) \quad \text{s.t. } P_0 \geq 0,\ Q \geq 0,\ R \geq 0 \qquad (5.5) \\ \tilde{\phi}(\mu, P_0, Q, R) &:= -2 \ln p_{\tilde{y}}(\tilde{y}; \mu, P_0, Q, R) + (T_f + 1) p \ln 2\pi \\ &= \ln |\tilde{P}_{\tilde{y}}| + (\tilde{y} - \mathcal{O}_{1,T_f+1}\mu)^{\mathrm{T}} \tilde{P}_{\tilde{y}}^{-1} (\tilde{y} - \mathcal{O}_{1,T_f+1}\mu)\end{aligned}$$

Problem (5.5) implicitly assumes that the inverse $\tilde{P}_{\tilde{y}}^{-1}$ exists, since it appears in the objective function.

ML problem (5.5) provides estimates of μ and P_0 in addition to estimates of Q and R. Estimates of μ and P_0 may be useful in some applications, but often they are nuisance parameters. Moreover, as T_f grows, the influence that μ and P_0 have on $\tilde{\phi}$ shrinks relative to that for Q and R. Therefore, we can expect the difference between $(\hat{Q}, \hat{R})$ from (5.3) and (5.5) to shrink as T_f grows. Finally, we remark that (5.5) includes general G and H matrices, but most EM-based covariance estimation methods assume $G = I_p$ and $H = I_n$.

Appendix 5.A derives the EM iteration equations that may be applied for solving (5.5). For this problem, the E step requires use of the Kalman smoother. The equations which arise from the M step for $(\hat{\mu}_{k+1}, \hat{P}_{0,k+1}, \hat{Q}_{k+1}, \hat{R}_{k+1})$ are explicit. In theory, EM can also be applied to solve (5.3), but for this problem the iterations cannot be obtained explicitly. Instead, the M step requires solution of a set of complicated nonlinear equations to get $(\hat{Q}_{k+1}, \hat{R}_{k+1})$. This detail is not included in Appendix 5.A, as we only discuss EM for solving (5.5). The interested reader may take it upon themself to derive the EM equations for (5.3).

5.3 Comments on the literature

Shumway and Stoffer (1982) were the first to use EM for the covariance estimation problem; estimation of A is included in their formulation as well. The details on estimating (μ, Σ) were not included in Shumway and Stoffer (1982), but given in a technical report of the same title (Shumway and Stoffer, 1981) that is no longer available according to David Stoffer's website.[1] However they are included in the

[1] `stat.pitt.edu/stoffer/dss.html`

authors' textbook (Shumway and Stoffer, 2017, Chapter 6).[2]

The covariance estimation ML problem has been extended in a number of directions. For instance, Ghahramani and Hinton (1996) use EM to provide ML estimates of C as well, and Ninness and Gibson (2001) and Gibson and Ninness (2005) use EM to solve a complete system identification problem. Bavdekar et al. (2011) develop an EM method which 1) solves an ML problem based on innovations instead of raw outputs, 2) uses an approximation of the likelihood function based on the assumption that the innovations are white to reduce the computational burden, and 3) uses the extended Kalman filter to apply the method to nonlinear systems. Ge and Kerrigan (2017b) develop an EM method for nonlinear systems which uses moving horizon or full information estimation instead of the extended Kalman filter, and allows for $G \neq I_n$ and $H \neq I_p$. Li and Badgwell (2014) develop an EM algorithm for linear systems which adds regularization terms to the problem to encourage sparse and low rank covariance estimates and allows for $G \neq I_n$ (but not for $H \neq I_p$). Zagrobelny and Rawlings (2014) (conference version: Zagrobelny and Rawlings (2015a)) implement several ML estimation algorithms, via both direct solution and EM, and compares the results and computation time with the ALS technique. We proceed similarly in Chapter 7, where we present numerical examples of ALS-based and ML-based algorithms.

Appendices

5.A Derivation of expectation maximization equations for covariance estimation

The EM algorithm presented in this appendix requires that P_0, Q, and R are invertible matrices. This assumption is reflected in several of the equations herein. Furthermore, we assume $G = I_n$ and $H = I_p$.

Assume that iteration k of the EM algorithm has been completed, i.e., $\hat{\mu}_k$, $\hat{P}_{0,k}$, $\hat{Q}_k$, and $\hat{R}_k$ are available. In this appendix, we show how to calculate the next EM iteration, i.e., we provide explicit expressions for $\hat{\mu}_{k+1}$, $\hat{P}_{0,k+1}$, $\hat{Q}_{k+1}$, and $\hat{R}_{k+1}$. First,

[2] David Stoffer notes that the Springer version of Shumway and Stoffer (2017) contains a number of errors. A corrected version may be downloaded at `stat.pitt.edu/stoffer/tsa4/tsa4.htm`.

we rearrange the full data likelihood function:

$$
\begin{aligned}
&p_{y_{0:T_f},x_{0:T_f}}(y_{0:T_f}, x_{0:T_f}) \\
&\quad = \frac{p_{x_0}(x_0)}{p_{x_0}(x_0)}\left(\prod_{j=0}^{T_f-1}\frac{p_{y_{0:j},x_{0:j}}(y_{0:j},x_{0:j})}{p_{y_{0:j},x_{0:j}}(y_{0:j},x_{0:j})}\frac{p_{y_{0:j},x_{0:j+1}}(y_{0:j},x_{0:j+1})}{p_{y_{0:j},x_{0:j+1}}(y_{0:j},x_{0:j+1})}\right)p_{y_{0:T_f},x_{0:T_f}}(y_{0:T_f},x_{0:T_f}) \\
&\quad = p_{x_0}(x_0)\frac{p_{y_0,x_0}(y_0,x_0)}{p_{x_0}(x_0)}\prod_{j=0}^{T_f-1}\frac{p_{y_{0:j},x_{0:j+1}}(y_{0:j},x_{0:j+1})}{p_{y_{0:j},x_{0:j}}(y_{0:j},x_{0:j})}\frac{p_{y_{0:j+1},x_{0:j+1}}(y_{0:j+1},x_{0:j+1})}{p_{y_{0:j},x_{0:j+1}}(y_{0:j},x_{0:j+1})} \\
&\quad = p_{x_0}(x_0)p_{y_0|x_0}(y_0 \mid x_0) \\
&\qquad\qquad \cdot \prod_{j=0}^{T_f-1} p_{x_{j+1}|y_{0:j},x_{0:j}}(x_{j+1} \mid y_{0:j},x_{0:j})p_{y_{j+1}|y_{0:j},x_{0:j+1}}(y_{j+1} \mid y_{0:j},x_{0:j+1}) \\
&\quad = p_{x_0}(x_0)p_{y_0|x_0}(y_0 \mid x_0)\prod_{j=0}^{T_f-1} p_{x_{j+1}|x_j}(x_{j+1} \mid x_j)p_{y_{j+1}|x_{j+1}}(y_{j+1} \mid x_{j+1}) \\
&\quad = p_{x_0}(x_0)\prod_{j=0}^{T_f-1} p_{x_{j+1}|x_j}(x_{j+1} \mid x_j)\prod_{j=0}^{T_f} p_{y_j|x_j}(y_j \mid x_j)
\end{aligned}
\tag{5.6}
$$

We have used the Markov property of the sequence $x_{0:T_f}$, which is the fact that the probability of a state conditioned on a set of prior states depends only on the most recent of the prior states (Anderson and Moore, 1979, page 18). Specifically, we have used

$$
\begin{aligned}
p_{x_{j+1}|y_{0:j},x_{0:j}}(x_{j+1} \mid y_{0:j},x_{0:j}) &= p_{x_{j+1}|x_j}(x_{j+1} \mid x_j) \\
p_{y_{j+1}|y_{0:j},x_{0:j+1}}(y_{j+1} \mid y_{0:j},x_{0:j+1}) &= p_{y_{j+1}|x_{j+1}}(y_{j+1} \mid x_{j+1})
\end{aligned}
$$

Now we explicitly give the three probability density functions needed in (5.6). First, $x_0 \sim \mathrm{N}(\mu, P_0)$, so

$$
p_{x_0}(x_0) = \frac{\exp[-\frac{1}{2}(x_0-\mu)^{\mathrm{T}}P_0^{-1}(x_0-\mu)]}{(2\pi)^{n/2}|P_0|^{1/2}} \tag{5.7}
$$

Next, $x_{j+1} = Ax_j + Bu_j + w_j$, so $x_{j+1} \mid x_j \sim \mathrm{N}(Ax_j + Bu_j, Q)$:

$$
\begin{aligned}
p_{x_{j+1}|x_j}(x_{j+1} \mid x_j) &= \frac{\exp(-\frac{1}{2}\bar{x}_j^{\mathrm{T}}Q^{-1}\bar{x}_j)}{(2\pi)^{n/2}|Q|^{1/2}} \\
\bar{x}_j &\coloneqq x_{j+1} - Ax_j - Bu_j
\end{aligned}
\tag{5.8}
$$

Similarly, $y_j = Cx_j + Du_j + v_j$, so $y_j \mid x_j \sim \mathrm{N}(Cx_j + Du_j, R)$:

$$p_{y_j|x_j}(y_j \mid x_j) = \frac{\exp(-\frac{1}{2}\bar{y}_j^{\mathrm{T}} R^{-1} \bar{y}_j)}{(2\pi)^{p/2}|R|^{1/2}} \qquad (5.9)$$
$$\bar{y}_j \coloneqq y_j - Cx_j - Du_j$$

Combining (5.6)–(5.9), we write the log likelihood function as

$$\begin{aligned}
&\ln p_{y_{0:T_f}, x_{0:T_f}}(y_{0:T_f}, x_{0:T_f}) \\
&\quad = \ln p_{x_0}(x_0) + \sum_{j=0}^{T_f-1} \ln p_{x_{j+1}|x_j}(x_{j+1} \mid x_j) + \sum_{j=0}^{T_f} \ln p_{y_j|x_j}(y_j \mid x_j) \\
&\quad = -\tfrac{n}{2}\ln 2\pi - \tfrac{1}{2}\ln|P_0| - \tfrac{1}{2}(x_0 - \mu)^{\mathrm{T}} P_0^{-1}(x_0 - \mu) \\
&\qquad + \sum_{j=0}^{T_f-1} [-\tfrac{n}{2}\ln 2\pi - \tfrac{1}{2}\ln|Q| - \tfrac{1}{2}\bar{x}_j^{\mathrm{T}} Q^{-1} \bar{x}_j] \\
&\qquad + \sum_{j=0}^{T_f} [-\tfrac{p}{2}\ln 2\pi - \tfrac{1}{2}\ln|R| - \tfrac{1}{2}\bar{y}_j^{\mathrm{T}} R^{-1} \bar{y}_j] \\
&\quad = -\tfrac{1}{2}\Big[(n+p)(T_f+1)\ln 2\pi + \ln|P_0| + (x_0 - \mu)^{\mathrm{T}} P_0^{-1}(x_0 - \mu) \\
&\qquad\qquad + T_f \ln|Q| + \sum_{j=0}^{T_f-1} \bar{x}_j^{\mathrm{T}} Q^{-1} \bar{x}_j + (T_f+1)\ln|R| + \sum_{j=0}^{T_f} \bar{y}_j^{\mathrm{T}} R^{-1} \bar{y}_j\Big]
\end{aligned}$$

The next step is to evaluate the conditional expectation

$$\mathrm{E}_{x_{0:T_f}|y_{0:T_f};\hat{\mu}_k,\hat{P}_{0,k},\hat{Q}_k,\hat{R}_k}[\ln p_{y_{0:T_f},x_{0:T_f}}(y_{0:T_f}, x_{0:T_f})] \qquad (5.10)$$

In order to take this conditional expectation, we need to know the conditional distribution $x_{0:T_f} \mid y_{0:T_f}; \hat{\mu}_k, \hat{P}_{0,k}, \hat{Q}_k, \hat{R}_k$. This distribution is obtained via the Kalman smoother:

$$\begin{bmatrix} x_{j+1} \\ x_j \end{bmatrix} \Bigg| \; y_{0:T_f} \,;\, \hat{\mu}_k, \hat{P}_{0,k}, \hat{Q}_k, \hat{R}_k \sim \mathrm{N}\left(\begin{bmatrix} \hat{x}^k_{j+1|T_f} \\ \hat{x}^k_{j|T_f} \end{bmatrix}, \begin{bmatrix} \Sigma^k_{j+1|T_f} & \Sigma^k_{j+1,j|T_f} \\ \Sigma^k_{j,j+1|T_f} & \Sigma^k_{j|T_f} \end{bmatrix} \right)$$

The expressions for the conditional mean $\hat{x}^k_{j|T_f}$, conditional covariance $\Sigma^k_{j|T_f}$, and conditional cross covariance $\Sigma^k_{j+1,j|T_f}$ are fairly involved, and so we omit a derivation.

We refer the reader to Anderson and Moore (1979, Chapter 7) and Yu et al. (2004) for derivations of the Kalman smoother equations.

Now we take the conditional expectation (5.10) in several steps. We make use of Searle (1971, Theorem 2.1, page 55) three times. First,

$$\begin{aligned}
&\mathrm{E}_{x_{0:T_f}|y_{0:T_f};\hat{\mu}_k,\hat{P}_{0,k},\hat{Q}_k,\hat{R}_k}[(x_0-\mu)^{\mathrm{T}}P_0^{-1}(x_0-\mu)] \\
&\quad = \operatorname{tr}(P_0^{-1}\Sigma^k_{0|T_f}) + (\hat{x}^k_{0|T_f}-\mu)^{\mathrm{T}}P_0^{-1}(\hat{x}^k_{0|T_f}-\mu) \\
&\quad = \operatorname{tr}(P_0^{-1}\Sigma^k_{0|T_f}) + \operatorname{tr}[(\hat{x}^k_{0|T_f}-\mu)^{\mathrm{T}}P_0^{-1}(\hat{x}^k_{0|T_f}-\mu)] \\
&\quad = \operatorname{tr}(P_0^{-1}\Sigma^k_{0|T_f}) + \operatorname{tr}[P_0^{-1}(\hat{x}^k_{0|T_f}-\mu)(\hat{x}^k_{0|T_f}-\mu)^{\mathrm{T}}] \\
&\quad = \operatorname{tr}\{P_0^{-1}[\Sigma^k_{0|T_f} + (\hat{x}^k_{0|T_f}-\mu)(\hat{x}^k_{0|T_f}-\mu)^{\mathrm{T}}]\}
\end{aligned}$$

Next,

$$\bar{x}_j = \begin{bmatrix} I_n & -A \end{bmatrix} \begin{bmatrix} x_{j+1} \\ x_j \end{bmatrix} - Bu_j$$

Hence

$$\begin{aligned}
&\bar{x}_j \mid y_{0:T_f};\, \hat{\mu}_k, \hat{P}_{0,k}, \hat{Q}_k, \hat{R}_k \\
&\quad \sim \mathrm{N}\left(\begin{bmatrix} I_n & -A \end{bmatrix} \begin{bmatrix} \hat{x}^k_{j+1|T_f} \\ \hat{x}^k_{j|T_f} \end{bmatrix} - Bu_j, \begin{bmatrix} I_n & -A \end{bmatrix} \begin{bmatrix} \Sigma^k_{j+1|T_f} & \Sigma^k_{j+1,j|T_f} \\ \Sigma^k_{j,j+1|T_f} & \Sigma^k_{j|T_f} \end{bmatrix} \begin{bmatrix} I_n \\ -A^{\mathrm{T}} \end{bmatrix} \right) \\
&\quad = \mathrm{N}(\underbrace{\hat{x}^k_{j+1|T_f} - A\hat{x}^k_{j|T_f} - Bu_j}_{=:\mu^k_{\bar{x},j}}, \underbrace{\Sigma^k_{j+1|T_f} - A\Sigma^k_{j,j+1|T_f} - \Sigma^k_{j+1,j|T_f}A^{\mathrm{T}} + A\Sigma^k_{j|T_f}A^{\mathrm{T}}}_{=:\Sigma^k_{\bar{x},j}})
\end{aligned}$$

Therefore

$$\begin{aligned}
\mathrm{E}_{x_{0:T_f}|y_{0:T_f};\hat{\mu}_k,\hat{P}_{0,k},\hat{Q}_k,\hat{R}_k}\left[\sum_{j=0}^{T_f-1} \bar{x}_j^{\mathrm{T}}Q^{-1}\bar{x}_j\right] &= \sum_{j=0}^{T_f-1} [\operatorname{tr}(Q^{-1}\Sigma^k_{\bar{x},j}) + (\mu^k_{\bar{x},j})^{\mathrm{T}}Q^{-1}\mu^k_{\bar{x},j}] \\
&= \sum_{j=0}^{T_f-1} [\operatorname{tr}(Q^{-1}\Sigma^k_{\bar{x},j}) + \operatorname{tr}\{(\mu^k_{\bar{x},j})^{\mathrm{T}}Q^{-1}\mu^k_{\bar{x},j}\}] \\
&= \sum_{j=0}^{T_f-1} [\operatorname{tr}(Q^{-1}\Sigma^k_{\bar{x},j}) + \operatorname{tr}\{Q^{-1}\mu^k_{\bar{x},j}(\mu^k_{\bar{x},j})^{\mathrm{T}}\}]
\end{aligned}$$

$$= \sum_{j=0}^{T_f-1} \operatorname{tr}\{Q^{-1}[\Sigma_{\tilde{x},j}^k + \mu_{\tilde{x},j}^k(\mu_{\tilde{x},j}^k)^{\mathrm{T}}]\}$$
$$= \operatorname{tr}\Bigg(Q^{-1} \underbrace{\sum_{j=0}^{T_f-1} [\Sigma_{\tilde{x},j}^k + \mu_{\tilde{x},j}^k(\mu_{\tilde{x},j}^k)^{\mathrm{T}}]}_{=:S_x^k}\Bigg)$$

Similarly,

$$\bar{y}_j \mid y_{0:T_f}\,;\, \hat{\mu}_k, \hat{P}_{0,k}, \hat{Q}_k, \hat{R}_k \sim \mathrm{N}(\underbrace{y_j - C\hat{x}_{j|T_f}^k - Du_j}_{=:\mu_{\bar{y},j}^k}, \underbrace{C\Sigma_{j|T_f}^k C^{\mathrm{T}}}_{\Sigma_{\bar{y},j}^k})$$

so

$$\mathrm{E}_{x_{0:T_f}|y_{0:T_f};\hat{\mu}_k,\hat{P}_{0,k},\hat{Q}_k,\hat{R}_k}\left[\sum_{j=0}^{T_f} \bar{y}_j^{\mathrm{T}} R^{-1} \bar{y}_j\right] = \sum_{j=0}^{T_f} [\operatorname{tr}(R^{-1}\Sigma_{\bar{y},j}^k) + (\mu_{\bar{y},j}^k)^{\mathrm{T}} R^{-1} \mu_{\bar{y},j}^k]$$
$$= \sum_{j=0}^{T_f} [\operatorname{tr}(R^{-1}\Sigma_{\bar{y},j}^k) + \operatorname{tr}\{(\mu_{\bar{y},j}^k)^{\mathrm{T}} R^{-1} \mu_{\bar{y},j}^k\}]$$
$$= \sum_{j=0}^{T_f} [\operatorname{tr}(R^{-1}\Sigma_{\bar{y},j}^k) + \operatorname{tr}\{R^{-1}\mu_{\bar{y},j}^k(\mu_{\bar{y},j}^k)^{\mathrm{T}}\}]$$
$$= \sum_{j=0}^{T_f} \operatorname{tr}\{R^{-1}[\Sigma_{\bar{y},j}^k + \mu_{\bar{y},j}^k(\mu_{\bar{y},j}^k)^{\mathrm{T}}]\}$$
$$= \operatorname{tr}\Bigg(R^{-1} \underbrace{\sum_{j=0}^{T_f} [\Sigma_{\bar{y},j}^k + \mu_{\bar{y},j}^k(\mu_{\bar{y},j}^k)^{\mathrm{T}}]}_{=:S_y^k}\Bigg)$$

We write the EM iteration as

$$\begin{aligned}
&(\hat{\mu}_{k+1}, \hat{P}_{0,k+1}, \hat{Q}_{k+1}, \hat{R}_{k+1}) \\
&\quad = \arg\max_{\mu,P_0,Q,R} \mathrm{E}_{x_{0:T_f}|y_{0:T_f};\hat{\mu}_k,\hat{P}_{0,k},\hat{Q}_k,\hat{R}_k}[\ln p_{y_{0:T_f},x_{0:T_f}}(y_{0:T_f}, x_{0:T_f})] \\
&\quad = \arg\min_{\mu,P_0,Q,R} \mathrm{E}_{x_{0:T_f}|y_{0:T_f};\hat{\mu}_k,\hat{P}_{0,k},\hat{Q}_k,\hat{R}_k}[-2\ln p_{y_{0:T_f},x_{0:T_f}}(y_{0:T_f}, x_{0:T_f}) \\
&\qquad\qquad - (n+p)(T_f+1)\ln 2\pi]
\end{aligned}$$

$$
\begin{aligned}
&= \arg\min_{\mu,P_0,Q,R} \mathrm{E}_{x_{0:T_f}|y_{0:T_f};\hat{\mu}_k,\hat{P}_{0,k},\hat{Q}_k,\hat{R}_k}\Big[\ln|P_0| + (x_0-\mu)^{\mathrm{T}}P_0^{-1}(x_0-\mu) \\
&\qquad\qquad + T_f\ln|Q| + \sum_{j=0}^{T_f-1}\bar{x}_j^{\mathrm{T}}Q^{-1}\bar{x}_j \\
&\qquad\qquad + (T_f+1)\ln|R| + \sum_{j=0}^{T_f}\bar{y}_j^{\mathrm{T}}Q^{-1}\bar{y}_j\Big] \\
&= \arg\min_{\mu,P_0,Q,R} J(\mu,P_0,Q,R) \qquad (5.11)
\end{aligned}
$$

where

$$
\begin{aligned}
J(\mu,P_0,Q,R) \coloneqq{}& \ln|P_0| + \mathrm{tr}\{P_0^{-1}[\Sigma^k_{0|T_f} + (\hat{x}^k_{0|T_f}-\mu)(\hat{x}^k_{0|T_f}-\mu)^{\mathrm{T}}]\} \\
&+ T_f\ln|Q| + \mathrm{tr}(Q^{-1}S_x^k) + (T_f+1)\ln|R| + \mathrm{tr}(R^{-1}S_y^k)
\end{aligned}
$$

We solve the optimization problem (5.11) by setting the differential of the objective function equal to zero, i.e., $\mathrm{d}J = 0$. Using several results on differentials from Magnus and Neudecker (2019, Chapters 8 and 9), we take the differential of each term of J:

$$
\begin{aligned}
\mathrm{d}\ln|P_0| &= \mathrm{tr}(P_0^{-1}\mathrm{d}P_0) \\
\mathrm{d}(T_f\ln|Q|) &= \mathrm{tr}(T_fQ^{-1}\mathrm{d}Q) \\
\mathrm{d}\,\mathrm{tr}(Q^{-1}S_x^k) &= \mathrm{tr}\,\mathrm{d}(Q^{-1}S_x^k) \\
&= \mathrm{tr}(-Q^{-1}(\mathrm{d}Q)Q^{-1}S_x^k) \\
\mathrm{d}[(T_f+1)\ln|R|] &= \mathrm{tr}[(T_f+1)R^{-1}\mathrm{d}R] \\
\mathrm{d}\,\mathrm{tr}(R^{-1}S_y^k) &= \mathrm{tr}\,\mathrm{d}(R^{-1}S_y^k) \\
&= \mathrm{tr}(-R^{-1}(\mathrm{d}R)R^{-1}S_y^k)
\end{aligned}
$$

$$
\begin{aligned}
&\mathrm{d}\,\mathrm{tr}\{P_0^{-1}[\Sigma^k_{0|T_f} + (\hat{x}^k_{0|T_f}-\mu)(\hat{x}^k_{0|T_f}-\mu)^{\mathrm{T}}]\} \\
&\quad= \mathrm{tr}\,\mathrm{d}\{P_0^{-1}[\Sigma^k_{0|T_f} + (\hat{x}^k_{0|T_f}-\mu)(\hat{x}^k_{0|T_f}-\mu)^{\mathrm{T}}]\} \\
&\quad= \mathrm{tr}\{\mathrm{d}P_0^{-1}[\Sigma^k_{0|T_f} + (\hat{x}^k_{0|T_f}-\mu)(\hat{x}^k_{0|T_f}-\mu)^{\mathrm{T}}] \\
&\qquad\quad + P_0^{-1}\mathrm{d}[\Sigma^k_{0|T_f} + (\hat{x}^k_{0|T_f}-\mu)(\hat{x}^k_{0|T_f}-\mu)^{\mathrm{T}}]\} \\
&\quad= \mathrm{tr}\{-P_0^{-1}(\mathrm{d}P_0)P_0^{-1}[\Sigma^k_{0|T_f} + (\hat{x}^k_{0|T_f}-\mu)(\hat{x}^k_{0|T_f}-\mu)^{\mathrm{T}}] \\
&\qquad\quad + P_0^{-1}[\mathrm{d}(\hat{x}^k_{0|T_f}-\mu)(\hat{x}^k_{0|T_f}-\mu)^{\mathrm{T}} + (\hat{x}^k_{0|T_f}-\mu)\mathrm{d}(\hat{x}^k_{0|T_f}-\mu)^{\mathrm{T}}]\}
\end{aligned}
$$

$$
\begin{aligned}
&= \operatorname{tr}\{-P_0^{-1}(\mathrm{d}P_0)P_0^{-1}[\Sigma_{0|T_f}^k + (\hat{x}_{0|T_f}^k - \mu)(\hat{x}_{0|T_f}^k - \mu)^{\mathrm{T}}] \\
&\qquad + P_0^{-1}[-\mathrm{d}\mu(\hat{x}_{0|T_f}^k - \mu)^{\mathrm{T}} - (\hat{x}_{0|T_f}^k - \mu)\mathrm{d}\mu^{\mathrm{T}}]\} \\
&= \operatorname{tr}[-P_0^{-1}\{(\mathrm{d}P_0)P_0^{-1}[\Sigma_{0|T_f}^k + (\hat{x}_{0|T_f}^k - \mu)(\hat{x}_{0|T_f}^k - \mu)^{\mathrm{T}}] \\
&\qquad + \mathrm{d}\mu(\hat{x}_{0|T_f}^k - \mu)^{\mathrm{T}} + (\hat{x}_{0|T_f}^k - \mu)\mathrm{d}\mu^{\mathrm{T}}\}]
\end{aligned}
$$

To make the $\mathrm{d}\mu$ terms equal zero, we must have

$$\mu = \hat{x}_{0:T_f}^k$$

Next, make the $\mathrm{d}P_0$ terms equal zero:

$$
\begin{aligned}
0 &= P_0^{-1}\mathrm{d}P_0 - P_0^{-1}\mathrm{d}(P_0)P_0^{-1}\Sigma_{0|T_f}^k \\
0 &= P_0^{-1}[\mathrm{d}P_0 - \mathrm{d}(P_0)P_0^{-1}\Sigma_{0|T_f}^k] \\
0 &= \mathrm{d}P_0 - \mathrm{d}(P_0)P_0^{-1}\Sigma_{0|T_f}^k \\
0 &= \mathrm{d}P_0[I_n - P_0^{-1}\Sigma_{0|T_f}^k] \\
0 &= I_n - P_0^{-1}\Sigma_{0|T_f}^k \\
I_n &= P_0^{-1}\Sigma_{0|T_f}^k \\
P_0 &= \Sigma_{0|T_f}^k
\end{aligned}
$$

Now make the $\mathrm{d}Q$ terms equal zero:

$$
\begin{aligned}
0 &= T_f Q^{-1}\mathrm{d}Q - Q^{-1}(\mathrm{d}Q)Q^{-1}S_x^k \\
0 &= Q^{-1}[T_f\mathrm{d}Q - (\mathrm{d}Q)Q^{-1}S_x^k] \\
0 &= T_f\mathrm{d}Q - (\mathrm{d}Q)Q^{-1}S_x^k \\
0 &= \mathrm{d}Q[T_f I_n - Q^{-1}S_x^k] \\
0 &= T_f I_n - Q^{-1}S_x^k \\
T_f I_n &= Q^{-1}S_x^k \\
Q &= \tfrac{1}{T_f}S_x^k
\end{aligned}
$$

Finally, make the $\mathrm{d}R$ terms equal zero:

$$
\begin{aligned}
0 &= (T_f + 1)R^{-1}\mathrm{d}R - R^{-1}(\mathrm{d}R)R^{-1}S_y^k \\
0 &= R^{-1}[(T_f + 1)\mathrm{d}R - (\mathrm{d}R)R^{-1}S_y^k]
\end{aligned}
$$

$$
\begin{aligned}
0 &= (T_f + 1)\mathrm{d}R - (\mathrm{d}R)R^{-1}S_y^k \\
0 &= \mathrm{d}R[(T_f + 1)I_p - R^{-1}S_y^k] \\
0 &= (T_f + 1)I_p - R^{-1}S_y^k \\
(T_f + 1)I_p &= R^{-1}S_y^k \\
R &= \tfrac{1}{T_f+1}S_y^k
\end{aligned}
$$

To summarize, the updated EM estimates are

$$
\begin{aligned}
\hat{\mu}_{k+1} &= \hat{x}_{0:T_f}^k \\
\hat{P}_{0,k+1} &= \Sigma_{0|T_f}^k \\
\hat{Q}_{k+1} &= \tfrac{1}{T_f}S_x^k \\
\hat{R}_{k+1} &= \tfrac{1}{T_f+1}S_y^k
\end{aligned}
$$

One cannot really argue with a mathematical theorem.

Stephen Hawking

CHAPTER 6

THEORETICAL COMPARISON OF COVARIANCE ESTIMATION METHODS

This chapter serves Objective 3 of the dissertation, which is to compare the ALS and ML approaches to covariance estimation. Specifically, this chapter discusses the theoretical properties which are guaranteed by various ALS-based and ML-based covariance estimation algorithms. There are many different metrics and properties which can be used as a basis for comparison when studying estimation algorithms. We focus our analysis on two of the most important and commonly used properties: unbiasedness and consistency. Definitions of unbiased and consistent may be found in Chapter 0 (Definitions 4 and 6, respectively).

The chapter is organized as follows:

- Section 6.1 presents and analyzes six ALS-based algorithms for covariance estimation, several of which have been enabled by the work on tractable calculation of the ALS weighting matrix from Chapter 4.
- Section 6.2 briefly discusses properties of ML-based algorithms.

The chapter also contains two appendices: 6.A and 6.B.

6.1 Properties of ALS-based algorithms

In this section, we discuss statistical properties (namely, bias and consistency) of ALS-based algorithms. First we discuss the properties of $\hat{b}$, which is necessary in order to analyze the properties of ALS estimates. We then offer six ALS-based algorithms and analysis thereof.

6.1.1 Properties of $\hat{\Lambda}_j$ and $\hat{b}$

Theorem 6.1. *For fixed $N \geq 0$, the estimate $\hat{b}$ is unbiased for b. Furthermore, if the noises (w_k) and (v_k) are jointly normally distributed as in (4.1), then $\hat{b}$ is consistent for b.*

The fact that $\hat{b}$ is unbiased has been noted since Odelson et al. (2006b), and was noted in Chapter 3, immediately beneath (3.5). Proof of consistency of $\hat{b}$ is a novel result as far as we are aware; see Appendix 6.B for the details.

Theorem 6.1 is equivalent to stating that $\hat{\Lambda}_{0:N-1}$ is unbiased and consistent, since $\hat{b}$ is the vectorized version of $\hat{\Lambda}_{0:N-1}$. If we consider the index of $\hat{\Lambda}$ to scale linearly with the number of data points instead of being fixed, we get a different result:

Theorem 6.2. *Assume that the noises (w_k) and (v_k) are jointly normally distributed as in (4.1). Then the estimate $\hat{\Lambda}_{T_f}$ is unbiased, but not consistent for Λ_{T_f}.*

Proof. We noted in Chapter 3, immediately beneath (3.4), that $\hat{\Lambda}_j$ is unbiased for arbitrary j. Using (2.21)–(2.24), we have

$$\begin{bmatrix} y_0 \\ y_{T_f} \end{bmatrix} \sim \mathrm{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Phi & (CA^{T_f-1}\Theta)^{\mathrm{T}} \\ CA^{T_f-1}\Theta & \Phi \end{bmatrix} \right)$$

Using Ghazal and Neudecker (2000, equation 3.3), we obtain

$$\begin{aligned} \operatorname{cov}[(\hat{\Lambda}_{T_f})_{\mathrm{s}}] = \operatorname{cov}[(y_{T_f} y_0^{\mathrm{T}})_{\mathrm{s}}] &= \operatorname{cov}(y_0 \otimes y_{T_f}) \\ &= (\Phi \otimes \Phi) + \mathscr{K}_p[CA^{T_f-1}\Theta \otimes (CA^{T_f-1}\Theta)^{\mathrm{T}}] \end{aligned}$$

which converges to $\Phi \otimes \Phi \neq 0$ as $T_f \to \infty$. However, because A is stable, $\Lambda_{T_f} = CA^{T_f-1}\Theta \to 0$ as $T_f \to \infty$. Therefore $\hat{\Lambda}_{T_f}$ is not consistent. ∎

6.1.2 ALS with arbitrary weight

Algorithm 1. Solve the ALS problem with arbitrary weighting.

Theorem 6.3. *Assume the ALS weighting, $\tilde{W}_\Omega$, satisfies $X_\Omega^{\mathrm{T}} \tilde{W}_\Omega X_\Omega > 0$ (i.e., the ALS objective function has a positive definite Hessian). Then Algorithm 1 with the unconstrained ALS problem yields an unbiased and consistent estimate of Ω.*

Proof. Unbiasedness was first established by Odelson et al. (2006b):

$$\begin{aligned}
\mathrm{E}[\hat{\Omega}] &= \mathrm{E}[(X_\Omega^{\mathrm{T}} \tilde{W}_\Omega X_\Omega)^+ X_\Omega^{\mathrm{T}} \tilde{W}_\Omega \hat{b}] \\
&= (X_\Omega^{\mathrm{T}} \tilde{W}_\Omega X_\Omega)^+ X_\Omega^{\mathrm{T}} \tilde{W}_\Omega \, \mathrm{E}[\hat{b}] \\
&= (X_\Omega^{\mathrm{T}} \tilde{W}_\Omega X_\Omega)^+ X_\Omega^{\mathrm{T}} \tilde{W}_\Omega X_\Omega \Omega \\
&= \Omega
\end{aligned}$$

The matrix Ω is a linear function of b, so consistency of $\hat{b}$ (Theorem 6.1) implies consistency of $\hat{\Omega}$ due to the continuous mapping theorem (van der Vaart, 1998, Theorem 2.3). ■

6.1.3 ALS with optimal weight

Algorithm 2. Solve the ALS problem using the optimal weighting $W_\Omega = (P_{\hat{b}} + X_\Omega X_\Omega^{\mathrm{T}})^+$.

Theorem 6.4. *Assume that $X_\Omega^{\mathrm{T}} W_\Omega X_\Omega > 0$ (i.e., the ALS objective function has a positive definite Hessian). Then Algorithm 2 with the unconstrained ALS problem yields an unbiased and consistent estimator of Ω.*

Proof of Theorem 6.4 is equivalent to that of Theorem 6.3, except W_Ω replaces $\tilde{W}_\Omega$. The assumption stated in Theorems 6.3 and 6.4, that $X_\Omega^{\mathrm{T}} W_\Omega X_\Omega > 0$, occurs if and only if X_Ω has LI columns. Recall that Section 3.2 discuss uniqueness conditions for each form of the ALS problem in more detail.

Algorithm 2 with the unconstrained ALS problem yields the minimum variance affine unbiased estimator of Ω from $\hat{b}$ (see Theorem A.2 in Appendix A), but it is not realizable in practice because W_Ω depends on the unknown true value of Ω. The purpose of considering Algorithm 2 is to use it as a benchmark to compare against other covariance estimation algorithms, both in terms of theoretical properties and empirically observed efficacy.

6.1.4 Direct estimation of the ALS weighting

Algorithm 3. Calculate an estimate of $P_{\hat{b}}$, denoted $\hat{P}_{\hat{b}}$, using (4.24), substituting $\hat{\Lambda}_j$ for Λ_j. Then solve the appropriate ALS problem using the weighting $\hat{W}_\Omega := (\hat{P}_{\hat{b}} + X_\Omega X_\Omega^{\mathrm{T}})^+$.

Algorithm 3 does not guarantee that either $\hat{P}_{\hat{b}}$ or $\hat{W}_\Omega$ are positive semidefinite, which is undesirable because the optimal values, $P_{\hat{b}}$ and W_Ω, do have this property. However, often the Hessian of the ALS objective function, $X_\Omega^{\mathrm{T}} \hat{W}_\Omega X_\Omega$, is positive semidefinite even if $\hat{W}_\Omega$ is indefinite; in this case, the ALS problem may be solved as usual. It is also possible that both $\hat{W}_\Omega$ and $X_\Omega^{\mathrm{T}} \hat{W}_\Omega X_\Omega$ can be indefinite, in which case the ALS objective function is unbounded from below. In this situation, one can use the equation from Theorem A.1 in Appendix A that gives the solution to the ALS problem when the Hessian is positive definite to get an estimate for Ω:

$$\hat{\Omega} = (X_\Omega^{\mathrm{T}} \hat{W}_\Omega X_\Omega)^+ X_\Omega^{\mathrm{T}} \hat{W}_\Omega \hat{b}$$

Such an estimate, however, will have no particular meaning relative to the ALS problem, and we make no claims as to any properties it might have. In contrast, consider the constrained ALS problem with an indefinite Hessian. It is possible that the ALS objective function can be bounded from below when restricted to the feasible region, but in this case there is no guarantee of a unique solution. In Chapter 7, we empirically study the frequency of occurrence for these scenarios (indefinite $\hat{W}_\Omega$ and indefinite $X_\Omega^{\mathrm{T}} \hat{W}_\Omega X_\Omega$).

Next we study statistical properties of $\hat{P}_{\hat{b}}$, $\hat{W}_\Omega$, and $\hat{\Omega}$ from Algorithm 3. We consider two cases: 1) $M < T_f$ is a fixed constant (i.e., barring nilpotency of A, (4.24) is an approximate expression for $P_{\hat{b}}$), and 2) $M \geq T_f$ (i.e., (4.24) is the exact expression for $P_{\hat{b}}$). In the first case, M fixed, $\hat{P}_{\hat{b}}$ is calculated from $\hat{\Lambda}_{0:M}$, and $\hat{\Lambda}_{0:M}$ is unbiased and consistent (Theorem 6.1). However, we have no guarantee that $\hat{P}_{\hat{b}}$ is either unbiased or consistent, since it is based on an approximation. In the second case, $M \geq T_f$, $\hat{P}_{\hat{b}}$ is calculated from $\hat{\Lambda}_{0:T_f}$, the latter of which is unbiased but not consistent (Theorem 6.2). Even though $\hat{\Lambda}_{0:T_f}$ is unbiased, because $P_{\hat{b}}$ is a nonlinear function of $\Lambda_{0:T_f}$, we have no guarantee that $\hat{P}_{\hat{b}}$ is unbiased.[1] Whether or not $\hat{P}_{\hat{b}}$ is consistent in this case remains an open question.

In any situation, we have no guarantee that $\hat{P}_{\hat{b}}$ is either unbiased or consistent. The weighting $\hat{W}_\Omega$ depends on $\hat{P}_{\hat{b}}$, and $\hat{\Omega}$ depends on $\hat{W}_\Omega$, so we also have no

[1] In general, for random variable X and nonlinear function $f(\cdot)$, $f(\mathrm{E}[X]) \neq \mathrm{E}[f(X)]$.

guarantee of unbiasedness or consistency for $\hat{W}_\Omega$ or $\hat{\Omega}$. Moreover, even if we could prove consistency of $\hat{P}_{\hat{b}}$, we could not then immediately conclude that either $\hat{W}_\Omega$ or $\hat{\Omega}$ is consistent—$\hat{W}_\Omega$ is not a continuous function of $\hat{P}_{\hat{b}}$ due to the presence of the Moore-Penrose inverse, and therefore the continuous mapping theorem cannot be invoked.

To be clear, we have not claimed that $\hat{P}_{\hat{b}}$, $\hat{W}_\Omega$, and $\hat{\Omega}$ are biased or inconsistent; we offer no guarantees in either direction. However, we can show that the estimate $\hat{P}_{\hat{b}}$ improves in some sense as $T_f \to \infty$. Consider M fixed and split (4.24) into components which do and do not depend on $\Lambda_{0:M}$:

$$P_{\hat{b}} = P_{\hat{b},\leq M} + P_{\hat{b},>M}$$

We offer the following two points:

1. The estimate $\hat{P}_{\hat{b}}$ is consistent for $P_{\hat{b},\leq M}$. This fact follows from the fact that $P_{\hat{b},\leq M}$ is a continuous function of $\Lambda_{0:M}$, consistency of $\hat{\Lambda}_{0:M}$ (Theorem 6.1), and the continuous mapping theorem.

2. For fixed T_f, $P_{\hat{b},>M} \to 0$ as $M \to T_f$, with $P_{\hat{b},>M} = 0$ for $M \geq T_f$.

In other words, $\hat{P}_{\hat{b}}$ is a consistent estimate of $P_{\hat{b},\leq M}$, and if M is sufficiently large, then the difference between $P_{\hat{b}}$ and $P_{\hat{b},\leq M}$, which is $P_{\hat{b},>M}$, is likely negligible for any practical purpose.

6.1.5 Iterative estimation of the ALS weighting

Algorithm 4.

1. Pick an initial weighting matrix, $\hat{W}_{\Omega,0}$, and use it to solve the ALS problem to give $\hat{\Omega}_0$.

2. Calculate $P_{\hat{b}}$ based on $\hat{\Omega}_0$ using (4.24); denote this quantity $\hat{P}_{\hat{b},0}$. Solve the ALS problem with the weighting matrix $\hat{W}_{\Omega,1} = (\hat{P}_{\hat{b},0} + X_\Omega X_\Omega^{\mathrm{T}})^+$ to give a new estimate, $\hat{\Omega}_1$.

3. Repeat the previous step until the sequence of estimates $(\hat{\Omega}_k)$ converges within a specified tolerance.

Algorithm 4 was first proposed by Rajamani and Rawlings (2009), where it is noted that the scheme is not guaranteed to converge. When using Algorithm 4, it is advisable to solve the constrained ALS problem at each iteration. Although it is possible to calculate Λ_j using (2.21)–(2.24), from (Q, R, S), $(Q, R, 0)$, or (Φ, Θ) which do not satisfy the respective semidefinite constraints, it is imprudent to expect that such Λ_j's would yield a reasonable value for $P_{\hat{b}}$ when substituted in (4.24).

Step 1 of Algorithm 4 is simply Algorithm 1. Thus, if one uses unconstrained ALS in step 1 and $X_\Omega^{\mathrm{T}} \hat{W}_{\Omega,0} X_\Omega > 0$, then $\hat{\Omega}_0$ is unbiased and consistent for Ω (Theorem 6.3). The matrix $\hat{P}_{\hat{b},0}$ is a continuous but nonlinear function of $\hat{\Omega}_0$, so by the continuous mapping theorem, $\hat{P}_{\hat{b},0}$ is consistent for $P_{\hat{b}}$, but there is no guarantee that it is unbiased. The matrix $\hat{W}_{\Omega,1}$ is not a continuous function of $\hat{P}_{\hat{b},0}$ due to the presence of the Moore-Penrose inverse, so the continuous mapping theorem cannot be invoked to guarantee that $\hat{W}_{\Omega,1}$ is consistent for W_Ω. In fact, for $k \geq 1$, none of $\hat{W}_{\Omega,k}$, $\hat{P}_{\hat{b},k}$, or $\hat{\Omega}_k$ are guaranteed to be either unbiased or consistent.

6.1.6 Previous methods to estimate the ALS weighting

Zagrobelny and Rawlings (2015b) was the first (and as far as we are aware, heretofore only) paper to study the issue of estimating the optimal ALS weighting matrix from data. The shortcomings from that paper are in part what inspired this dissertation, so next we present the results from Zagrobelny and Rawlings (2015b) as a means to compare with this work.

Zagrobelny and Rawlings (2015b) use a different estimate for b than $\hat{b}$ in (3.5):

$$\hat{b}_Z \coloneqq \left(\frac{1}{\tilde{n}+1} \sum_{i=0}^{\tilde{n}} c_i y_i^{\mathrm{T}} \right)_{\mathrm{ss}} = \mathscr{D}_{\tilde{p},p}^{+} \mathcal{J}(\hat{P}_c)_{\mathrm{s}}$$

$$\hat{P}_c \coloneqq \frac{1}{\tilde{n}+1} \sum_{i=0}^{\tilde{n}} c_i c_i^{\mathrm{T}}$$

Like $\hat{b}$, the vector $\hat{b}_Z$ is unbiased for b. The difference is that $\hat{b}_Z$ does not use c_i, $\tilde{n}+1 \leq i \leq T_f$, which are the c_i's containing zero blocks. The covariance of $\hat{b}_Z$ is

$$P_{\hat{b}_Z} \coloneqq \operatorname{cov}(\hat{b}_Z) = \mathscr{D}_{\tilde{p},p}^{+} \mathcal{J} P_{\hat{P}_c} \mathcal{J}^{\mathrm{T}} (\mathscr{D}_{\tilde{p},p}^{+})^{\mathrm{T}} \tag{6.1}$$

$$P_{\hat{P}_c} \coloneqq \operatorname{cov}[(\hat{P}_c)_{\mathrm{s}}]$$

Thus estimating $P_{\hat{b}_Z}$ is akin to estimating $P_{\hat{P}_c}$. Zagrobelny and Rawlings (2015b) cite Magnus and Neudecker (1979, Corollary 4.2) to note that under the assumption

that c_i and c_j, $i \neq j$, are independent and identically distributed zero mean normal random variables, then

$$P_{\hat{P}_c} = \frac{1}{\tilde{n}+1}(I_{\tilde{p}^2} + \mathscr{K}_{\tilde{p}})(P_c \otimes P_c) \tag{6.2}$$
$$P_c := \operatorname{cov}(c_i)$$

The c_i's meet all of these assumptions except that they are not independent. Recall that if, based on the stability of A, we approximate $A^k = 0$ for all $k \geq M$, then c_i and c_{i+k} are uncorrelated for $k \geq N + M$. For the normal distribution, uncorrelated and independence are equivalent (Graham and Rawlings, 2013, Example 4.9). Therefore, Zagrobelny and Rawlings (2015b) proposes separating the c_i's into groups of approximately independent columns. Assuming $N_s := \frac{\tilde{n}+1}{N+M}$ is an integer, we have

$$Y_i = \begin{bmatrix} c_i & c_{N+M+i} & \cdots & c_{(N_s-1)(N+M)+i} \end{bmatrix}, \quad 0 \leq i \leq N+M-1$$

Let $\hat{P}_{c,i}$ be the sample covariance of the columns of Y_i:

$$\hat{P}_{c,i} = \frac{1}{N_s}\sum_{j=0}^{N_s-1} c_{(N+M)j+i}c_{(N+M)j+i}^{\mathrm{T}}, \quad 0 \leq i \leq N+M-1$$

Zagrobelny and Rawlings (2015b) propose averaging the $\hat{P}_{c,i}$ matrices to get an estimate for $\hat{P}_c$ and substituting into (6.2):

$$\hat{P}_{c,Z} = \frac{1}{N+M}\sum_{i=0}^{N+M-1} \hat{P}_{c,i}$$
$$\hat{P}_{\hat{P}_c}^{Z,1} = \frac{1}{\tilde{n}+1}(I_{\tilde{p}^2} + \mathscr{K}_{\tilde{p}})(\hat{P}_{c,Z} \otimes \hat{P}_{c,Z}) \tag{6.3}$$

It is easy to verify that if N_s is an integer, then $\hat{P}_{c,Z} = \hat{P}_c$. Thus, dividing the c_i's into approximately independent groups is not a suitable workaround for the fact that (6.2) is an inappropriate equation upon which to base an estimate of $P_{\hat{P}_c}$.

The above strategy is actually slightly different than what is implemented in the ALS software package[2] that was used for the numerical examples in Zagrobelny and Rawlings (2015b). What the ALS package does instead is

$$\hat{P}_{\hat{P}_{c,i}} = \frac{1}{N_s}(I_{\tilde{p}^2} + \mathscr{K}_{\tilde{p}})(\hat{P}_{c,i} \otimes \hat{P}_{c,i})$$
$$\hat{P}_{\hat{P}_c}^{Z,2} = \frac{1}{N+M}\sum_{i=0}^{N+M-1} \hat{P}_{\hat{P}_{c,i}} \tag{6.4}$$

[2]Available: `bitbucket.org/rawlings-group/als`.

The matrix $\hat{P}_{\hat{P}_c,i}$ would be a reasonable estimate of $P_{\hat{P}_c}$ if only the columns of Y_i were used to calculate $\hat{P}_c$. Since this situation is not realized, $\hat{P}^{Z,2}_{\hat{P}_c}$ is theoretically suspect as an estimate of $P_{\hat{P}_c}$. Relative to $\hat{P}^{Z,1}_{\hat{P}_c}$, the estimate $\hat{P}^{Z,2}_{\hat{P}_c}$ tends to overestimate in magnitude the scalar variances and covariances which constitute $P_{\hat{P}_c}$. The reason is clear upon examination: combining equations, $\hat{P}^{Z,1}_{\hat{P}_c}$ has a leading scalar factor of $1/(\tilde{n}+1)^3$ whereas $\hat{P}^{Z,2}_{\hat{P}_c}$ has a leading scalar factor of $1/[(\tilde{n}+1)N_s^2]$, and the denominator of the former is generally significantly larger than that of the latter.

We now present two ALS algorithms based on (6.3) and (6.4):

Algorithm 5. Calculate an estimate of $P_{\hat{b}_Z}$, denoted $\hat{P}_{\hat{b}_Z}$, using (6.1), substituting $\hat{P}^{Z,1}_{\hat{P}_c}$ for $P_{\hat{P}_c}$. Then solve the appropriate ALS problem using the weighting matrix $\hat{W}^{Z,1}_{\Omega} := (\hat{P}_{\hat{b}_Z} + X_\Omega X_\Omega^{\mathrm{T}})^+$.

Algorithm 6. Calculate an estimate of $P_{\hat{b}_Z}$, denoted $\hat{P}_{\hat{b}_Z}$, using (6.1), substituting $\hat{P}^{Z,2}_{\hat{P}_c}$ for $P_{\hat{P}_c}$. Then solve the appropriate ALS problem using the weighting matrix $\hat{W}^{Z,2}_{\Omega} := (\hat{P}_{\hat{b}_Z} + X_\Omega X_\Omega^{\mathrm{T}})^+$.

Because of the unrealistic assumptions upon which (6.3) and (6.4) are based, neither Algorithm 5 nor Algorithm 6 guarantees that any of $\hat{P}^{Z,k}_{\hat{P}_c}$, $\hat{P}_{\hat{b}_Z}$, $\hat{W}_\Omega$ and $\hat{\Omega}$ are unbiased or consistent estimators. One advantage of Algorithms 5 and 6 relative to Algorithm 3 is that they guarantee that $\hat{P}_{\hat{b}_Z}$ and $\hat{W}_\Omega$ are positive semidefinite.

6.2 Properties of ML-based algorithms

ML estimation methods are broadly popular for many applications partly because they have desirable statistical properties. The theoretical properties of ML estimation have been well studied, but such analysis is relatively technical and we do not wish to bog down this dissertation with the details. Instead, we simply state three important facts regarding ML estimates:

1. Under general assumptions, ML estimates are consistent (Newey and McFadden, 1994, Theorem 2.5).

2. Under stronger assumptions, ML estimates also are asymptotically normal and efficient (an estimator is efficient if it achieves the Cramer-Rao lower bound (Kay, 1993, Chapter 3)) (Newey and McFadden, 1994, Theorem 3.3).

3. ML estimates are not guaranteed to be unbiased, although in some cases they are.

Appendices

6.A Miscellaneous results

Lemma 6.1. *For scalars $x \geq 0$ and $y \geq 0$, $\sqrt{x+y} \leq \sqrt{x} + \sqrt{y}$.*

Proof. We write

$$x + y \leq x + 2\sqrt{xy} + y = (\sqrt{x} + \sqrt{y})^2$$

Because the square root function is nondecreasing, we may take the square root of both sides of this inequality to obtain the desired result. ■

Lemma 6.2. *The Frobenius norm of a partitioned matrix satisfies*

$$\left|\!\left|\!\left|\begin{bmatrix} A & B \\ C & D \end{bmatrix}\right|\!\right|\!\right|_F = \left|\!\left|\!\left|\begin{bmatrix} |||A|||_F & |||B|||_F \\ |||C|||_F & |||D|||_F \end{bmatrix}\right|\!\right|\!\right|_F \leq |||A|||_F + |||B|||_F + |||C|||_F + |||D|||_F$$

Proof.

$$\begin{aligned}
\left|\!\left|\!\left|\begin{bmatrix} A & B \\ C & D \end{bmatrix}\right|\!\right|\!\right|_F &= \sqrt{\sum_{i,j} |A_{i,j}|^2 + \sum_{i,j} |B_{i,j}|^2 + \sum_{i,j} |C_{i,j}|^2 + \sum_{i,j} |D_{i,j}|^2} \\
&= \sqrt{|||A|||_F^2 + |||B|||_F^2 + |||C|||_F^2 + |||D|||_F^2} \\
&= \left|\!\left|\!\left|\begin{bmatrix} |||A|||_F & |||B|||_F \\ |||C|||_F & |||D|||_F \end{bmatrix}\right|\!\right|\!\right|_F \\
&\leq \sqrt{\sum_{i,j} |A_{i,j}|^2} + \sqrt{\sum_{i,j} |B_{i,j}|^2} + \sqrt{\sum_{i,j} |C_{i,j}|^2} + \sqrt{\sum_{i,j} |D_{i,j}|^2} \\
&= |||A|||_F + |||B|||_F + |||C|||_F + |||D|||_F
\end{aligned}$$

where we have used Lemma 6.1. ■

Lemma 6.3. *For matrices A and B,*

$$|||A \otimes B|||_F = |||A|||_F |||B|||_F$$

Proof. The scalar $\sigma_{A\otimes B}$ is a singular value of $A\otimes B$ if and only if $\sigma_{A\otimes B} = \sigma_A\sigma_B$ where σ_A and σ_B are singular values of A and B, respectively (Graham and Rawlings, 2013, equation 1.26, page 67). Thus, we obtain

$$|||A\otimes B|||_F = \sqrt{\sum_i \sigma^2_{A\otimes B,i}} = \sqrt{\sum_{i,j}\sigma^2_{A,i}\sigma^2_{B,j}} = \sqrt{\Big(\sum_i \sigma^2_{A,i}\Big)\Big(\sum_i \sigma^2_{B,i}\Big)}$$
$$= \sqrt{\sum_i \sigma^2_{A,i}}\sqrt{\sum_i \sigma^2_{B,i}} = |||A|||_F|||B|||_F$$

where we have used the fact that for matrix X, $|||X|||_F = \sqrt{\operatorname{tr} X^{\mathrm{T}}X} = \sqrt{\sum_i \sigma^2_{X,i}}$. ■

6.B Proof of Theorem 6.1

Proof. First, note that stability of A implies there exist scalar constants c and γ with $c > 0$ and $0 < \gamma < 1$ such that $|(A^k)_{ij}| \leq c\gamma^k$ for all natural numbers k, where $(A^k)_{ij}$ denotes each individual scalar element of the matrix A^k (Horn and Johnson, 2013, Corollary 5.6.13, page 349). Therefore

$$|||A^k|||_F = \|(A^k)_{\mathrm{s}}\|_2 \leq \|(A^k)_{\mathrm{s}}\|_1 = \sum_{i,j}|(A^k)_{i,j}| \leq \underbrace{n^2 c}_{=:\alpha_A}\gamma^k$$

where we have made use of the fact that $\|x\|_2 \leq \|x\|_1$ for any vector x (Golub and Van Loan, 2013, equation 2.2.5, page 69). Next, $\Lambda_j = CA^{j-1}\Theta$ for $j \geq 1$. Therefore

$$\begin{aligned}
|||\Lambda_j|||_F &= \|(\Lambda_j)_{\mathrm{s}}\|_2 \\
&= \|(CA^{j-1}\Theta)_{\mathrm{s}}\|_2 \\
&= \|(\Theta^{\mathrm{T}}\otimes C)(A^{j-1})_{\mathrm{s}}\|_2 \\
&\leq |||\Theta^{\mathrm{T}}\otimes C|||_2\|(A^{j-1})_{\mathrm{s}}\|_2 \\
&= |||\Theta^{\mathrm{T}}\otimes C|||_2|||A^{j-1}|||_F \\
&\leq \underbrace{|||\Theta^{\mathrm{T}}\otimes C|||_2\alpha_A}_{=:\alpha_\Lambda}\gamma^{j-1}, \quad j \geq 1
\end{aligned}$$

where the first inequality follows from Golub and Van Loan (2013, inequality 2.3.2, page 71). Next, for $j \geq N$, we have

$$\begin{aligned}
|||P_{c,j,0}|||_F &\leq N \sum_{i=-N+1}^{N-1} |||\Lambda_{j+i}|||_F \\
&\leq N\alpha_\Lambda \sum_{i=-N+1}^{N-1} \gamma^{j+i-1} \\
&\leq N\alpha_\Lambda \sum_{i=-N+1}^{N-1} \gamma^{j-N} \\
&\leq N(2N-1)\alpha_\Lambda \gamma^{j-N} \\
&= \underbrace{N(2N-1)\alpha_\Lambda \gamma^{-N}}_{=:\tilde{\alpha}_P} \gamma^j, \quad j \geq N
\end{aligned}$$

where we have used Lemma 6.2 and the fact that γ^x is a nonincreasing function of x. Then for $j \geq 0$, we have

$$\begin{aligned}
|||P_{c,j,0}|||_F &\leq \alpha_P \gamma^j, \quad j \geq 0 \\
\alpha_P &:= \max(\tilde{\alpha}_P, \max_{0 \leq j \leq N-1} |||P_{c,j,0}|||_F \gamma^{-j})
\end{aligned}$$

Next, for $j \geq 0$,

$$\begin{aligned}
|||P_{\otimes,j,0}|||_F &= |||(I_{\tilde{p}} + \mathscr{K}_{\tilde{p}})(P_{c,j,0} \otimes P_{c,j,0})|||_F \\
&\leq |||I_{\tilde{p}} + \mathscr{K}_{\tilde{p}}|||_F |||P_{c,j,0} \otimes P_{c,j,0}|||_F \\
&= |||I_{\tilde{p}} + \mathscr{K}_{\tilde{p}}|||_F |||P_{c,j,0}|||_F^2 \\
&\leq \underbrace{|||I_{\tilde{p}} + \mathscr{K}_{\tilde{p}}|||_F \alpha_P^2}_{=:\alpha_\otimes} \gamma^{2j}
\end{aligned}$$

where we have used Lemma 6.3 and the fact that the Frobenius norm is submultiplicative, i.e., $|||AB|||_F \leq |||A|||_F |||B|||_F$ (Horn and Johnson, 2013, pages 341–342).[3] Also, note that $|||P_{\otimes,j,k}|||_F \leq |||P_{\otimes,j,0}|||_F$ for all $k \geq 0$ and $j \geq 0$, since $P_{\otimes,j,k}$ and $P_{\otimes,j,0}$ are the same except that the latter may have some block rows and columns of

[3] Horn and Johnson technically define the Frobenius norm only for square matrices, but the method they use to derive submultiplicativity extends readily to the case with nonsquare matrices.

zeros, depending on k and j. Continuing, we obtain

$$
\begin{aligned}
|||P_{c_\otimes}|||_F &= \left|\left|\left|\begin{bmatrix} P_{\otimes,0,0} & P^{\mathrm{T}}_{\otimes,1,0} & \cdots & P^{\mathrm{T}}_{\otimes,\tilde{n},0} \\ P_{\otimes,1,0} & P_{\otimes,0,1} & \cdots & P^{\mathrm{T}}_{\otimes,\tilde{n}-1,1} \\ \vdots & \vdots & \ddots & \vdots \\ P_{\otimes,\tilde{n},0} & P_{\otimes,\tilde{n}-1,1} & \cdots & P_{\otimes,0,\tilde{n}} \end{bmatrix}\right|\right|\right|_F \\
&= \left|\left|\left|\begin{bmatrix} |||P_{\otimes,0,0}|||_F & |||P^{\mathrm{T}}_{\otimes,1,0}|||_F & \cdots & |||P^{\mathrm{T}}_{\otimes,\tilde{n},0}|||_F \\ |||P_{\otimes,1,0}|||_F & |||P_{\otimes,0,1}|||_F & \cdots & |||P^{\mathrm{T}}_{\otimes,\tilde{n}-1,1}|||_F \\ \vdots & \vdots & \ddots & \vdots \\ |||P_{\otimes,\tilde{n},0}|||_F & |||P_{\otimes,\tilde{n}-1,1}|||_F & \cdots & |||P_{\otimes,0,\tilde{n}}|||_F \end{bmatrix}\right|\right|\right|_F \\
&\leq \left|\left|\left|\begin{bmatrix} |||P_{\otimes,0,0}|||_F & |||P^{\mathrm{T}}_{\otimes,1,0}|||_F & \cdots & |||P^{\mathrm{T}}_{\otimes,\tilde{n},0}|||_F \\ |||P_{\otimes,1,0}|||_F & |||P_{\otimes,0,0}|||_F & \cdots & |||P^{\mathrm{T}}_{\otimes,\tilde{n}-1,0}|||_F \\ \vdots & \vdots & \ddots & \vdots \\ |||P_{\otimes,\tilde{n},0}|||_F & |||P_{\otimes,\tilde{n}-1,0}|||_F & \cdots & |||P_{\otimes,0,0}|||_F \end{bmatrix}\right|\right|\right|_F \\
&\leq \left|\left|\left|\alpha_\otimes \begin{bmatrix} \gamma^0 & \gamma^2 & \cdots & \gamma^{2\tilde{n}} \\ \gamma^2 & \gamma^0 & \cdots & \gamma^{2\tilde{n}-2} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma^{2\tilde{n}} & \gamma^{2\tilde{n}-2} & \cdots & \gamma^0 \end{bmatrix}\right|\right|\right|_F \\
&= \alpha_\otimes \left|\left|\left|\begin{bmatrix} \gamma^0 & \gamma^2 & \cdots & \gamma^{2\tilde{n}} \\ \gamma^2 & \gamma^0 & \cdots & \gamma^{2\tilde{n}-2} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma^{2\tilde{n}} & \gamma^{2\tilde{n}-2} & \cdots & \gamma^0 \end{bmatrix}\right|\right|\right|_F \\
&\leq \alpha_\otimes \sqrt{\tilde{n}+1} \left|\left|\left|\begin{bmatrix} \gamma^0 & \gamma^2 & \cdots & \gamma^{2\tilde{n}} \\ \gamma^2 & \gamma^0 & \cdots & \gamma^{2\tilde{n}-2} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma^{2\tilde{n}} & \gamma^{2\tilde{n}-2} & \cdots & \gamma^0 \end{bmatrix}\right|\right|\right|_\infty \\
&\leq 2\alpha_\otimes \sqrt{\tilde{n}+1} \sum_{i=0}^{\tilde{n}} \gamma^{2i} \\
&\leq 2\alpha_\otimes \sqrt{\tilde{n}+1} \sum_{i=0}^{\infty} \gamma^{2i}
\end{aligned}
$$

$$= 2\alpha_{\otimes}\sqrt{\tilde{n}+1}\sum_{i=0}^{\infty}(\gamma^2)^i$$
$$= \frac{2\alpha_{\otimes}\sqrt{\tilde{n}+1}}{1-\gamma^2}$$

where we have used Lemma 6.2 and the well known infinite geometric series summation formula (Rudin, 1976, Theorem 3.26, page 61). Next, observe

$$|||\mathcal{F}|||_F = \sqrt{p}\left|\left|\left|\begin{bmatrix}\frac{1}{T_f+1} & & \\ & \frac{1}{T_f} & \\ & & \ddots \\ & & & \frac{1}{\tilde{n}+1}\end{bmatrix}\right|\right|\right|_F = \sqrt{\sum_{i=1}^{N}\frac{p}{(\tilde{n}+i)^2}} \leq \sqrt{\frac{pN}{(\tilde{n}+1)^2}} = \frac{\sqrt{\tilde{p}}}{\tilde{n}+1}$$
$$|||\mathcal{G}|||_F = \sqrt{p}$$
$$|||\mathscr{D}^+_{\tilde{p},p}|||_F = \sqrt{p\tilde{p} - \tfrac{3}{4}p(p-1)}$$

The expression for $|||\mathscr{D}^+_{\tilde{p},p}|||_F$ follows because each row of $\mathscr{D}^+_{\tilde{p},p}$ has either one 1 entry or two 0.5 entries, and the rest zeros. Therefore

$$\begin{aligned}
|||\mathcal{H}|||_F &= |||\mathscr{D}^+_{\tilde{p},p}(\mathcal{G}\otimes\mathcal{F})|||_F \\
&\leq |||\mathscr{D}^+_{\tilde{p},p}|||_F|||\mathcal{G}\otimes\mathcal{F}|||_F \\
&\leq |||\mathscr{D}^+_{\tilde{p},p}|||_F|||\mathcal{G}|||_F|||\mathcal{F}|||_F \\
&\leq \sqrt{p\tilde{p} - \tfrac{3}{4}p(p-1)}\sqrt{p}\frac{\sqrt{\tilde{p}}}{\tilde{n}+1} \\
&= \underbrace{\sqrt{(p\tilde{p})^2 - \tfrac{3}{4}p^2(p-1)\tilde{p}}}_{=:\alpha_{\mathcal{H}}}\frac{1}{\tilde{n}+1}
\end{aligned}$$

where again we have used submultiplicativity of the Frobenius norm and Lemma 6.3. Next:

$$|||\mathcal{E}_j|||_F \leq |||I_{\tilde{p}}|||_F = \sqrt{\tilde{p}}, \quad 0 \leq j \leq N-1$$
$$|||\tilde{\mathcal{E}}|||_F = \left|\left|\left|\sum_{j=0}^{N-1}(\mathcal{E}_j\otimes\mathcal{E}_j)\right|\right|\right|_F \leq \sum_{j=0}^{N-1}|||\mathcal{E}_j\otimes\mathcal{E}_j|||_F = \sum_{j=0}^{N-1}|||\mathcal{E}_j|||_F^2 \leq \sum_{j=0}^{N-1}\tilde{p} = N\tilde{p}$$

where we have used the triangle inequality (Golub and Van Loan, 2013, page 71) and Lemma 6.3. Therefore

$$\begin{aligned}
|||\mathcal{U}|||_F &= \left|\left|\left|\begin{bmatrix}(\imath_{\tilde{n}}^{\mathrm{T}} \otimes I_{\tilde{p}^2}) & \tilde{\mathcal{E}}\end{bmatrix}\right|\right|\right|_F \\
&\leq |||\imath_{\tilde{n}}^{\mathrm{T}} \otimes I_{\tilde{p}^2}|||_F + |||\tilde{\mathcal{E}}|||_F \\
&= |||\imath_{\tilde{n}}^{\mathrm{T}}|||_F |||I_{\tilde{p}^2}|||_F + |||\tilde{\mathcal{E}}|||_F \\
&\leq \sqrt{\tilde{n}}\,\tilde{p} + N\tilde{p} \\
&= \tilde{p}(\sqrt{\tilde{n}} + N)
\end{aligned}$$

where we have used Lemmas 6.2 and 6.3. Then

$$\begin{aligned}
|||P_{\hat{b}}|||_F &= |||\mathcal{H} P_a \mathcal{H}^{\mathrm{T}}|||_F \\
&= |||\mathcal{H}\mathcal{U} P_{c_\otimes} \mathcal{U}^{\mathrm{T}} \mathcal{H}^{\mathrm{T}}|||_F \\
&\leq |||\mathcal{H}|||_F^2 |||\mathcal{U}|||_F^2 |||P_{c_\otimes}|||_F \\
&\leq \frac{\alpha_{\mathcal{H}}^2}{(\tilde{n}+1)^2} \tilde{p}^2 (\tilde{n} + 2N\sqrt{\tilde{n}} + N^2) \frac{2\alpha_\otimes \sqrt{\tilde{n}+1}}{1-\gamma^2} \\
&= \frac{2\alpha_{\mathcal{H}}^2 \alpha_\otimes \tilde{p}^2}{1-\gamma^2} \frac{\sqrt{\tilde{n}+1}}{(\tilde{n}+1)^2} (\tilde{n} + 2N\sqrt{\tilde{n}} + N^2) \\
&= \frac{2\alpha_{\mathcal{H}}^2 \alpha_\otimes \tilde{p}^2}{1-\gamma^2} (\tilde{n}+1)^{-3/2} (\tilde{n} + 2N\sqrt{\tilde{n}} + N^2) \\
&\leq \frac{2\alpha_{\mathcal{H}}^2 \alpha_\otimes \tilde{p}^2}{1-\gamma^2} (\tilde{n}+1)^{-3/2} [(\tilde{n}+1) + 2N\sqrt{\tilde{n}+1} + N^2] \\
&= \frac{2\alpha_{\mathcal{H}}^2 \alpha_\otimes \tilde{p}^2}{1-\gamma^2} [(\tilde{n}+1)^{-1/2} + 2N(\tilde{n}+1)^{-1} + N^2(\tilde{n}+1)^{-3/2}] \\
&\leq \frac{2\alpha_{\mathcal{H}}^2 \alpha_\otimes \tilde{p}^2}{1-\gamma^2} [(\tilde{n}+1)^{-1/2} + 2N(\tilde{n}+1)^{-1/2} + N^2(\tilde{n}+1)^{-1/2}] \\
&= \frac{2\alpha_{\mathcal{H}}^2 \alpha_\otimes \tilde{p}^2}{1-\gamma^2} (1 + 2N + N^2)(\tilde{n}+1)^{-1/2} \\
&= \underbrace{\frac{2\alpha_{\mathcal{H}}^2 \alpha_\otimes \tilde{p}^2 (N+1)^2}{1-\gamma^2}}_{=:\alpha} (\tilde{n}+1)^{-1/2}
\end{aligned}$$

where we have used (4.4) and (4.12) and submultiplicativity of the Frobenius norm. The constant α does not depend on T_f, and $\tilde{n} = T_f - N + 1$, so we conclude that

$|||P_{\hat{b}}|||_F \to 0$ as $T_f \to \infty$. Finally, we have

$$\begin{aligned}
\lim_{T_f\to\infty} \mathrm{E}[\|\hat{b}-b\|_2^2] &= \lim_{T_f\to\infty} \mathrm{E}[(\hat{b}-b)^{\mathrm{T}}(\hat{b}-b)] \\
&= \lim_{T_f\to\infty} \mathrm{E}[\mathrm{tr}\{(\hat{b}-b)(\hat{b}-b)^{\mathrm{T}}\}] \\
&= \lim_{T_f\to\infty} \mathrm{tr}\,\mathrm{E}[(\hat{b}-b)(\hat{b}-b)^{\mathrm{T}}] \\
&= \lim_{T_f\to\infty} \mathrm{tr}\,P_{\hat{b}} \\
&= \mathrm{tr}\Big(\lim_{T_f\to\infty} P_{\hat{b}}\Big) \\
&= 0
\end{aligned}$$

where we have used the facts that $\mathrm{tr}\,AB = \mathrm{tr}\,BA$ and that, since it is a linear operation, trace commutes with the expectation and limit operations. We have demonstrated that $\hat{b}$ converges to b in mean square, which implies convergence in probability, and hence consistency of $\hat{b}$ (Gubner, 2006, chapters 13 and 14). ■

The best way out is always through.

Robert Frost

CHAPTER 7

EMPIRICAL COMPARISON OF COVARIANCE ESTIMATION METHODS

This chapter continues the work of the previous chapter in serving Objective 3 of the dissertation to compare the ALS and ML approaches to covariance estimation. This chapter shifts the focus to numerical examples, using simulated data to study the relative efficacy of ALS-based and ML-based estimation methods.

The chapter is organized as follows:

- Section 7.1 lists 17 covariance estimation methods, applies each method to 100 different simulated data sets, plots the results, and discusses the relative quality of the covariance estimates obtained from each method.
- Section 7.2 discusses the effect of the number of data points, N_d, on the computation time required for ALS-based and ML-based covariance estimation methods.

7.1 Quality of covariance matrix estimates

In this section, we use the following model:

$$A = \begin{bmatrix} 0.7 & 0.3 \\ -0.2 & -0.8 \end{bmatrix} \qquad C = I_2 \qquad Q = \begin{bmatrix} 10 & 0 \\ 0 & 0.1 \end{bmatrix} \qquad R = \begin{bmatrix} 1 & 0 \\ 0 & 0.01 \end{bmatrix}$$
$$G = I_2 \qquad H = I_2$$

We simulated 100 different data sets from this model, each with $N_d = 500$.

7.1.1 List of covariance estimation methods used

For each data set, we used 17 different methods to estimate Q and R, assuming $S = 0$:

Method 1. Estimate Q and R as the sample covariances of $w_{0:T_f-1}$ and $v_{0:T_f}$, respectively.

Method 2. Algorithm 1: ALS, unconstrained, $W_{QR} = I$.

Method 3. Algorithm 1: ALS, constrained, $W_{QR} = I$.

Method 4. Algorithm 2: ALS, unconstrained, W_{QR} = true optimal weighting.

Method 5. Algorithm 2: ALS, constrained, W_{QR} = true optimal weighting.

Method 6. Algorithm 3: ALS, unconstrained, $W_{QR} = \hat{W}_{QR}$, $M = 20$.

Method 7. Algorithm 3: ALS, constrained, $W_{QR} = \hat{W}_{QR}$, $M = 20$.

Method 8. Algorithm 5: ALS, unconstrained, $W_{QR} = \hat{W}_{QR}^{Z,1}$, $M = 20$.

Method 9. Algorithm 5: ALS, constrained, $W_{QR} = \hat{W}_{QR}^{Z,1}$, $M = 20$.

Method 10. Algorithm 6: ALS, unconstrained, $W_{QR} = \hat{W}_{QR}^{Z,2}$, $M = 20$.

Method 11. Algorithm 6: ALS, constrained, $W_{QR} = \hat{W}_{QR}^{Z,2}$, $M = 20$.

Method 12. Algorithm 4: iterative ALS, constrained, 10 iterations, $\hat{W}_{QR,0} = I$, $M = T_f$.

Method 13. Algorithm 4: iterative ALS, constrained, 10 iterations, $\hat{W}_{QR,0} = I$, $M = T_f$, assume Q and R are diagonal.

Method 14. ML problem (5.5), solved by performing 100 EM iterations, with initial guesses $\hat{Q}_0 = \hat{R}_0 = I_2$.

Method 15. ML problem (5.5) with Q and R diagonal, solved by performing 100 EM iterations, with initial guesses $\hat{Q}_0 = \hat{R}_0 = I_2$.[1]

Method 16. ML problem (5.3), solved via Matlab's `fmincon` function with the interior point algorithm.

Method 17. ML problem (5.3) with Q and R diagonal and sparse $P_{\tilde{y}}$ ($M = 20$), solved via Matlab's `fmincon` function with the interior point algorithm.

For all ALS-based methods, $N = 10$ was used as the lag parameter. Method 1 is not realizable in practice because the sequences $w_{0:T_f-1}$ and $v_{0:T_f}$ are unknown, and Methods 4 and 5 are not realizable in practice because the true optimal weighting matrix depends on the true values of Q and R, which are unknown. However, these methods can be implemented with simulated data sets, since $w_{0:T_f-1}$, $v_{0:T_f}$, Q, and R are known in this case. We include Methods 1, 4, and 5 in our study to provide a basis for comparison with the other methods.

The simulations were conducted in Matlab version R2020a on a quad core Intel Core i5-7500T with 16 GB RAM. The open source Covest package[2] was used to solve all ALS problems and to calculate the smoothed state estimates as required in each step of the EM algorithm in Methods 14 and 15.

7.1.2 Observed sample bias and sample variance in simulation results

Figures 7.1–7.4 plot the estimates from the simulations. Recalling the results from Section 6.1, the only methods which yield theoretically unbiased estimates are sample covariance (Method 1), unconstrained ALS with identity weighting (Method 2), and unconstrained ALS with optimal weighting (Method 4). Of these methods, only Method 2 is realizable in practice. However, several other methods exhibit a small amount of bias in our simulations: constrained ALS with true optimal weighting

[1] Recall that the EM iteration equations presented in Appendix 5.A do not assume that Q and R are diagonal. It is a straightforward exercise to derive the EM iteration equations under the assumption that Q and R are diagonal; we leave this task as an exercise for the interested reader.

[2] Available: `bitbucket.org/rawlings-group/covest`

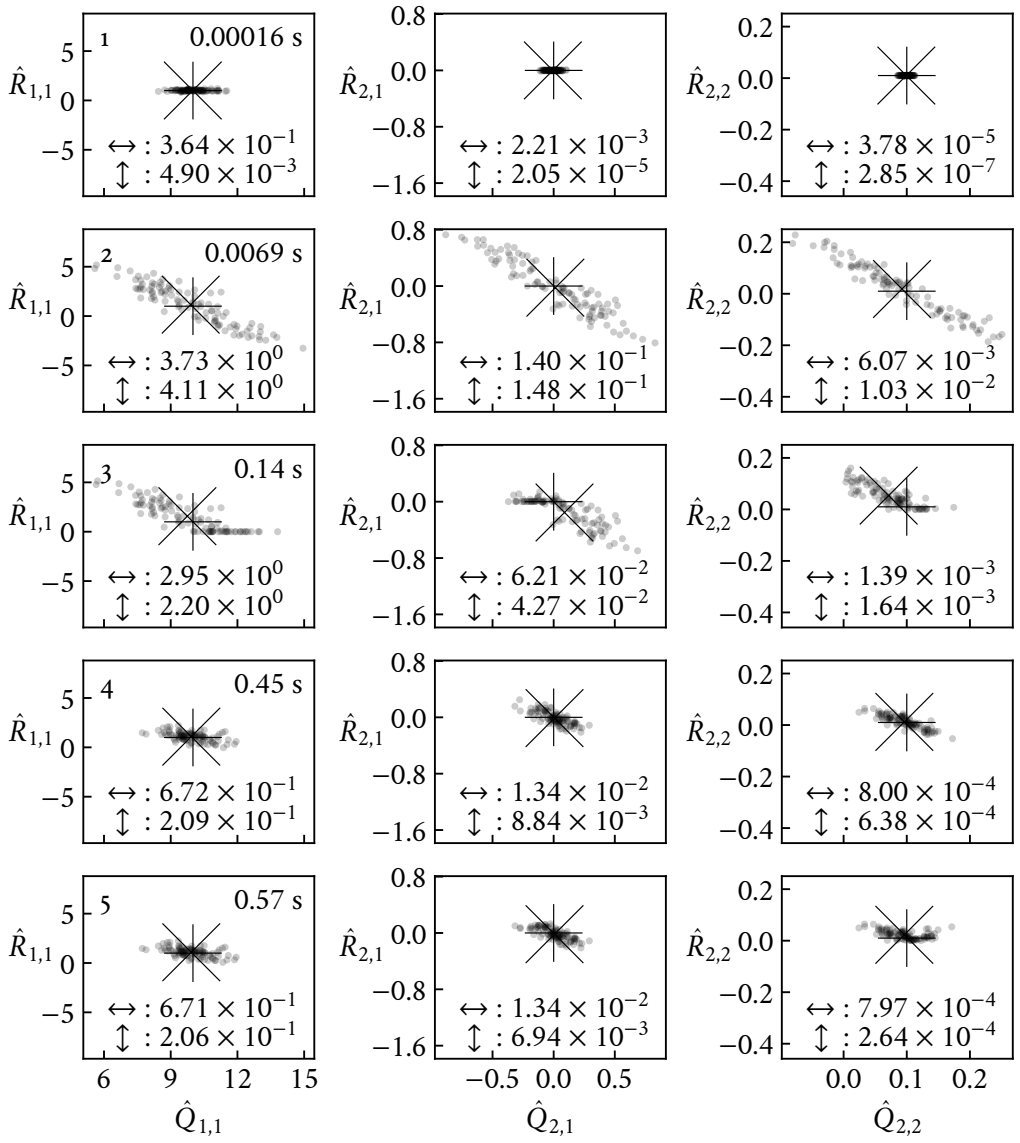

Figure 7.1: Covariance estimation simulation results, 1 of 4, ALS methods. In the left column, the number in the upper left is method for the row and the number in the upper right is the average time for the method. In each plot, the center of the + is the true value and the center of the × is the sample mean. The numbers next to ↔ and ↕ represent the sample variance of the variables plotted along the x and y axes, respectively. Each column of plots has the same axes limits for ease of comparison.

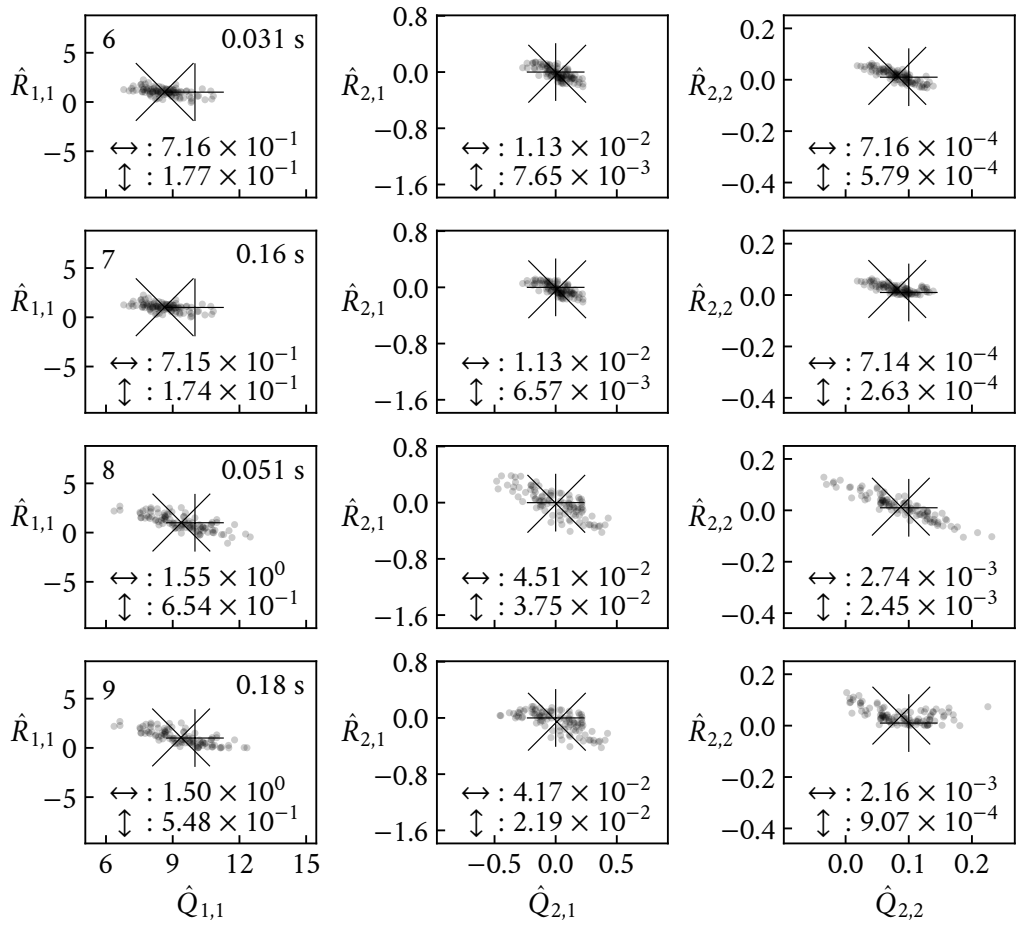

Figure 7.2: Covariance estimation simulation results, 2 of 4, ALS methods.

(Method 5), constrained ALS with iterative weighting (Methods 12 and 13), and all of the ML-based estimation methods (Methods 14–17). Of these methods, all except Method 5 are realizable.

The realizable methods which yield covariance estimates with the smallest sample variance are the ML estimation methods (Methods 14–17), but some of the ALS methods (specifically, Methods 4–7, 12, and 13) yield estimates with only slightly larger sample variances. Moreover, Methods 13, 15, and 17 demonstrate that assuming Q and R diagonal yields lower variance estimates of the diagonal elements of Q and R, both for ALS-based and ML-based estimation methods.

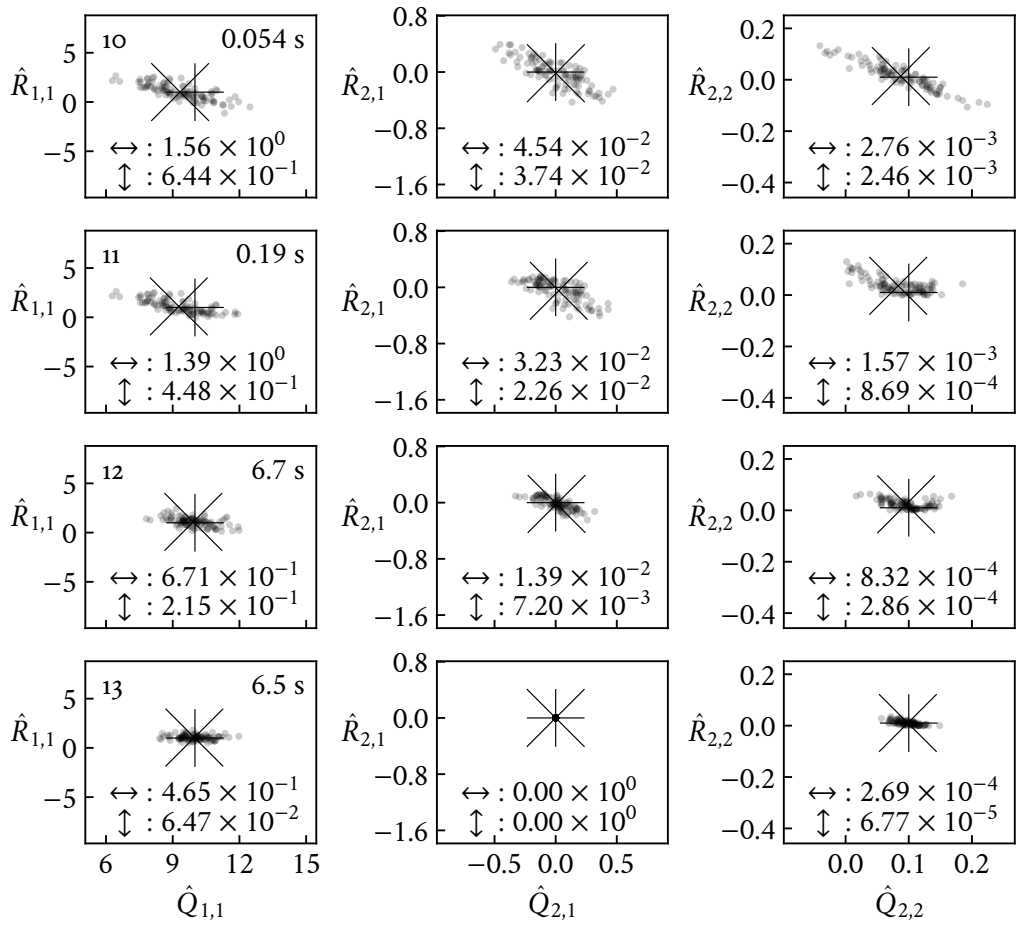

Figure 7.3: Covariance estimation simulation results, 3 of 4, ALS methods.

7.1.3 Effect of lag parameter N for ALS-based methods

For the ALS methods (Methods 2–13), we solved all of the problems using $N = 5$ and $N = 20$ in addition to $N = 10$ to study the effect of changing N. We have not included plots of these results to preserve organization in this dissertation, but we offer the following comments on the results:

- For Methods 2–5, 12, and 13, N has a minor effect on the covariance estimates.
- For Methods 6 and 7, $N \in \{5, 10\}$ has a minor effect. However, with $N = 20$, these methods occasionally produce outliers. For example, for one particular

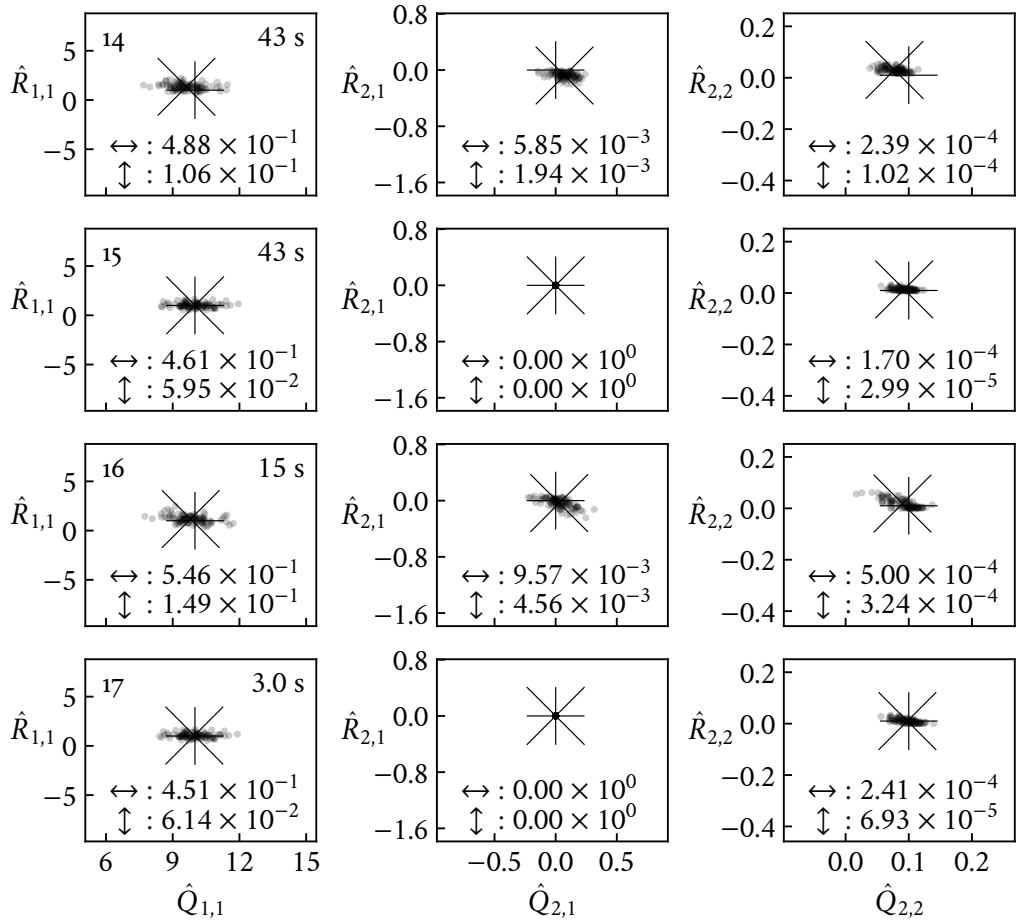

Figure 7.4: Covariance estimation simulation results, 4 of 4, ML methods.

data set, Method 6 with $N = 20$ gives

$$\hat{Q} = \begin{bmatrix} -22.1 & 4.20 \\ 4.20 & -0.84 \end{bmatrix} \qquad \hat{R} = \begin{bmatrix} 21.0 & -2.56 \\ -2.56 & 0.59 \end{bmatrix}$$

For such outliers, the matrix $(X_{QR}^{\mathrm{T}} \hat{W}_{QR} X_{QR})^{-1} X_{QR}^{\mathrm{T}} \hat{W}_{QR}$ tends to have a relatively high condition number compared to other instances. Furthermore, with $N \in \{5, 10\}$, for each data set, Methods 6 and 7 yield a positive semidefinite weighting $\hat{W}_{QR}$ and Hessian $X_{QR}^{\mathrm{T}} \hat{W}_{QR} X_{QR}$. However, with $N = 20$, 24 of the 100 problems have an indefinite weighting, and two of these also have an indefinite Hessian.

- For Methods 8–11, increasing N decreases the variance and increases the bias of the covariance estimates.

7.1.4 Comments on ML-based methods

Among the ML-based estimation methods, EM (Methods 14 and 15) gives noticeably different estimates than direct ML (Methods 16 and 17). The difference is not drastic, but perhaps larger than one might expect. The reason for the discrepancy is unclear, but we offer three possible explanations:

1. ML problems (5.3) and (5.5) indeed do have noticeably different solutions for some example problems.
2. Not enough EM iterations were performed in solving (5.5), resulting in slightly suboptimal solutions.
3. Matlab's `fmincon` yields suboptimal solutions to (5.3) for some problems. We noticed this issue with `fmincon` optimization algorithms other than interior point, namely SQP and active set, but it is unclear if, and if so to what extent, it occurs with interior point.

7.1.5 Convergence rates for iterative ALS and EM

Figure 7.5 demonstrates the convergence rates for Methods 12 and 14. We offered no guarantee of convergence for ALS with iterative weighting (Method 12), but in Figure 7.5, linear convergence is observed for every sample problem. In comparison, the convergence rate of ML estimation via EM (Method 14) is quite slow. The Q and R diagonal versions, (Methods 13 and 15), are not included in Figure 7.5, but similar convergence rates are achieved.

7.2 Effect of N_d on computation time

Figures 7.1–7.4 indicate that in the simulations from the previous section, ALS-based methods generally require less computation time than ML-based methods. The exceptions are Methods 12 and 13, which take significantly longer than the other ALS-based methods because they solve 10 least squares problems in succession,

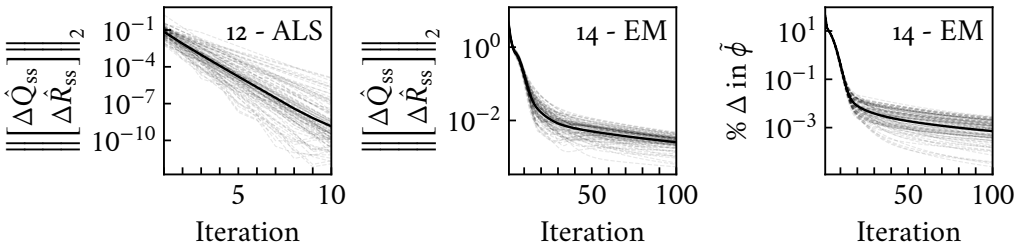

Figure 7.5: Convergence rates for iterative ALS (Method 12) and ML estimation via EM (Method 14). The first and second plots show the norm of the difference in the estimate at each iteration for iterative ALS and EM, respectively. The third plot shows the percent change in the ML objective function $\tilde{\phi}$ at each EM iteration. Each gray dotted line represents a particular simulation and the solid black line is the geometric mean of all simulations.

whereas all of the other ALS-based methods solve only one least squares problem. However, these observations do not tell the full story. Our next task is to study the effect of the number of data points, N_d, on the computation time for ALS and ML-based estimation methods.

We generated a random, stable, 10 state system:

$$\texttt{A = randn(10); A = A/max(abs(eig(A)))/2;} \tag{7.1}$$
$$C = I_{10} \qquad Q = I_{10} \qquad R = I_{10} \qquad G = I_{10} \qquad H = I_{10}$$

Equation (7.1) is the line of Matlab code that was used to create A. We varied N_d and studied the computation time of ALS and ML estimation as follows:

- We solved four different ALS problems: with and without constraints and with and without the diagonal Q and R assumption. We calculated the weighting using the true values of Λ_j to avoid generating an ALS problem with an indefinite Hessian. We took $M = 20$ because using larger M has a negligible effect on the solution, but it makes the weighting calculation take significantly longer.
- We performed one EM iteration.
- As a proxy for directly solving the ML problem (5.3), we evaluated the ML objective function ϕ once, both with and without assuming sparse $P_{\tilde{y}}$ and that Q and R are diagonal.

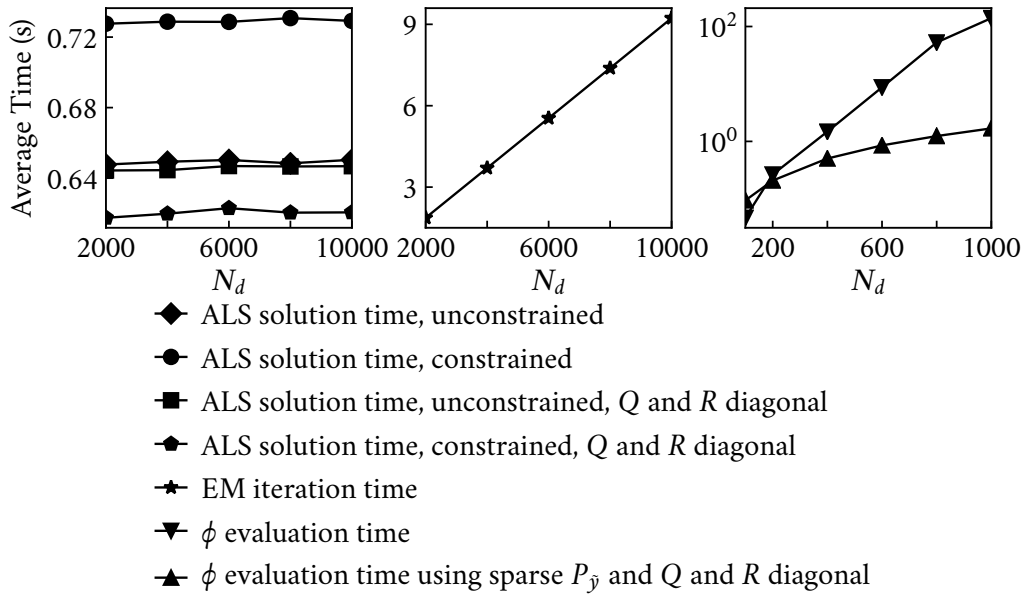

Figure 7.6: Effect of N_d on computation time for various covariance estimation methods. Note the log scale on the third plot.

The average computation times for 100 simulations are plotted in Figure 7.6. The ALS computation time is not dependent on N_d. This result is expected, as the dimension of the ALS problem does not depend on N_d. As N_d increases, evaluating $\hat{b}$ takes longer, but the time required for this calculation is negligible relative to that of solving the least squares problem.

The EM iteration computation time scales linearly with N_d. This result is expected, as the majority of the computational burden associated with an EM iteration is running the Kalman smoother on the data, a process for which the computation time scales linearly with N_d. Assuming Q and R diagonal does not change the Kalman smoother, so under this assumption the EM computation time is unaffected.

There are two factors which contribute to the total computation time required for EM: the time of each individual iteration and the number of iterations needed to converge. We have not precisely studied how the required number of iterations scales with N_d (although we have already shown in Figure 7.5 that the rate of convergence for EM is relatively slow). Nevertheless, observing the linear dependence of the individual EM iteration time on N_d is sufficient to conclude that the time required for the entire EM algorithm must scale at least linearly with N_d (in reality,

it probably scales somewhat superlinearly).

Similarly, the computation time for direct ML solution depends on both the number of iterations required from the optimization algorithm and on the time for each individual iteration. The number of optimizer iterations required is rather variable (for example, in the previous section, Method 16 required 49 iterations on average with a standard deviation of 11), but it generally increases as N_d increases. As with EM, we have not carefully studied the N_d vs. iterations required relationship, but each iteration requires multiple evaluations of ϕ, and Figure 7.5 shows that the time for just one such evaluation increases rapidly with N_d. With the sparsity and diagonality assumptions, the scaling is slightly superlinear, but without them it is nearly exponential. Note that for $N_d = 200$, the smallest value tested, it takes more time to evaluate ϕ with the diagonality and sparsity assumptions compared to without them. This phenomenon occurs because $P_{\tilde{y}}$ is still relatively dense for this number of data points, so treating it as sparse actually is a handicap computationally.

We tried using the solution from ALS with iterative weighting (Method 12) as a warm start initial guess to decrease the computation time of the ML estimation methods, via both direct solution and EM. This strategy does moderately reduce the ML computation time, but it is not strong enough of an effect to significantly alter the undesirable scaling between N_d and ML computation time.

A good hockey player plays where the puck is. A great hockey player plays where the puck is going to be.

Wayne Gretzky

CHAPTER 8

CONCLUSION: SUMMARY AND SUGGESTIONS FOR FUTURE WORK

In this chapter, we restate the three guiding research objectives from Chapter 1. We evaluate the extent to which each objective is achieved by summarizing relevant contributions and achievements from the dissertation. Furthermore, we suggest avenues for possible research related to each objective. We also include a section specifically dedicated to covariance estimation for linear time-varying and nonlinear systems, as this potential direction for future research is common across all three objectives.

8.1 Identifiability of covariance matrices (Q, R, S) and Kalman gain matrices (K, L)

Objective 1. Obtain precise results regarding the identifiability of the covariance matrices (Q, R, S), the Kalman predictor gain matrix K, and the Kalman filter gain matrix L.

8.1.1 Summary of contributions

The identifiability analysis in this dissertation in Chapters 2 and 3 is based on the autocovariance least squares (ALS) technique for covariance estimation, which can only be applied to an output sequence which is wide-sense stationary. However, the outputs are not necessarily wide-sense stationary if the system has inputs and/or if the matrix A is not stable. Therefore, we conduct our analysis in two parallel tracks: for the process system without inputs and matrix A stable, (1.3) and (1.4), and also for the $\mathcal{K}$-error system, (2.10) and (2.11). The $\mathcal{K}$-error system generates the $\mathcal{K}$-innovations sequence, $\mathbb{z}_{0:T_f}$, which is wide-sense stationary.

Most previous work on covariance estimation considers the matrices Q and R and assumes $S = 0$. In addition to $(Q, R, 0)$, we also study ALS estimation of (Q, R, S) as well as for a different set of parameters: (Φ, Θ) for the process system and (Φ, Θ) for the $\mathcal{K}$-error system. We obtain the following precise results for parameter identifiability under the assumptions $G = I_n$ and $H = I_p$:

- Matrices (Q, R, S) are not uniquely identifiable.
- Matrices $(Q, R, 0)$ are uniquely identifiable if and only if C has LI columns (which is equivalent to saying that every state of the system model is measured) and A is invertible.
- Matrices (Φ, Θ) and (Φ, Θ) are uniquely identifiable if and only if (A, C) is observable.
- The Kalman predictor gain K can be uniquely determined from (Φ, Θ) or (Φ, Θ). Therefore, K is uniquely identifiable if (A, C) is observable.
- The Kalman filter gain L cannot be uniquely determined from (Φ, Θ) or (Φ, Θ), generally speaking. However, if A is invertible and it is known (or assumed) that $S = 0$, then L and R can be uniquely determined from (Φ, Θ) or (Φ, Θ).

Furthermore, we demonstrate that if $G \neq I_n$ and/or $H \neq I_p$, then these identifiability properties can change; for example, if $G \neq I_n$, then in some cases (Q, R, S) is uniquely identifiable. Additionally, if the objective is to estimate K, then, under general assumptions, solving any of the (Q, R, S), (Φ, Θ), or (Φ, Θ) versions of the ALS problem will result in the same estimate for K (see Theorems 3.3 and 3.4 in Section 3.3).

To be concise, we can conceptualize our results as a sort of hierarchy of identifiability

$$(\Phi, \Theta) = (\Phi, \Theta) > (Q, R, 0) > (Q, R, S)$$
$$K > L$$

in which the parameters on the left have the least restrictive requirements in order to be identified uniquely, and the parameters on the right the most restrictive requirements.

Although our identifiability analysis is couched in terms of only one particular correlation method, ALS, the conclusions we reach apply to other correlation methods as well. Specifically, if the ALS problem does not have a unique solution, then it is impossible to form any correlation method which will yield a unique estimate. Let us explain. Recall that the relationship between the autocovariance and parameters $\Omega \in \{(Q, R, S), (Q, R, 0), (\Phi, \Theta), (\Phi, \Theta)\}$ is linear, i.e., $b = X_\Omega \Omega$ (or $\mathbb{b} = \mathbb{X}_\Omega \Omega$), and that the ALS problem has a unique solution if and only if X_Ω ($\mathbb{X}_\Omega$) has LI columns. However, it is a fundamental consequence of the linear relationship $b = X_\Omega \Omega$ ($\mathbb{b} = \mathbb{X}_\Omega \Omega$) that if X_Ω ($\mathbb{X}_\Omega$) has LD columns, then b ($\mathbb{b}$) cannot uniquely determine Ω. Therefore, in this case, there is no correlation method that yields a unique estimate of Ω.

8.1.2 Suggestions for future work

Control system performance monitoring

The favorable identifiability properties of (Φ, Θ) and (Φ, Θ) is perhaps the most important practical reason for considering these parameters. For example, consider the application of controller performance monitoring, briefly discussed in Chapter 1. With this task, it is desirable to track parameters that can be identified uniquely, so that when a monitoring application detects that the parameters have changed, the control engineer knows that it must be because the underlying noise/disturbance characteristics of the system have changed. For example, Zagrobelny et al. (2013) describe a monitoring scheme which depends on knowledge of the covariance matrices Q and R, assuming $S = 0$. However, if for a given application $(Q, R, 0)$ cannot be identified uniquely, then this monitoring scheme might prove ineffectual. A monitoring scheme based on identification of (Φ, Θ) or (Φ, Θ) is likely to be more useful, so one potential direction for future work is to develop such a monitoring scheme.

The relationship between ALS and ML estimation

The results regarding identifiability via ALS from this dissertation also have implications on identifiability via the maximum likelihood methods. We offer the following conjecture:

Conjecture 8.1. *Assume that the noises (w_k) and (v_k) are jointly normally distributed, as in (4.1). Then the ML problem to estimate Ω has a unique solution if and only if X_Ω with $N = T_f + 1$ has LI columns.*

The "only if" part of Conjecture 8.1 is straightforward to prove. Recall from (5.2) that the probability density of $y_{0:T_f}$ is parameterized by $P_{\tilde{y}}$. From (4.7), $P_{\tilde{y}}$ is a function of $\Lambda_{0:T_f}$; from (3.1), Ψ is a matrix of $\Lambda_{0:T_f}$ vertically stacked; and from (3.2), $\Psi_{\text{ss}} = X_\Omega \Omega$. Therefore, the probability density function $p_{y_{0:T_f}}(y_{0:T_f})$ can be written as a function ℓ of $X_\Omega \Omega$, i.e., $p_{y_{0:T_f}}(y_{0:T_f}; P_{\tilde{y}}) = \ell(X_\Omega \Omega)$. So, if $\hat{\Omega}_1$ solves the ML problem, then for $\hat{\Omega}_2 = X_\Omega^+ X_\Omega (\hat{\Omega}_1 - q) + q$ where q is an arbitrary vector, we have $X_\Omega \hat{\Omega}_1 = X_\Omega \hat{\Omega}_2$, so $\hat{\Omega}_2$ also solves the ML problem.

Results from Zagrobelny and Rawlings (2014) (conference version: Zagrobelny and Rawlings (2015a)) and Rajamani (2007, Chapter 6) suggest that the "if" part of Conjecture 8.1 is true as well, though these references assume $G = I_n$ and $H = I_p$ and only consider $\Omega = (Q, R, 0)$. Future work should seek a unified and rigorous treatment of the connection between the ALS and ML problems, including potentially a proof of Conjecture 8.1.

The identifiability question more broadly

In our presentation of ALS in this dissertation, we consider only the second moment (i.e., the autocovariance) of the outputs (or $\mathcal{K}$-innovations) and we make no assumptions on the distribution of the process and measurement noises. The most common choice in practice is to assume a normal distribution for these noises (which we have done in Chapters 4–7). For the normal distribution, the first and second moments (i.e., the mean and covariance) completely parameterize the distribution and determine all higher moments. Thus, studying higher moments of the outputs (or $\mathcal{K}$-innovations) cannot possibly yield further information about the noise covariances than can be gleaned from studying the output autocovariance. However, it is conceivable that if the noises are not normally distributed, then in some cases it may be possible to extract additional information from higher moments of the outputs (or $\mathcal{K}$-innovations). For example, suppose that for a given

system we have $G = I_n$ and $H = I_p$, and therefore (Q, R, S) cannot be uniquely identified from the autocovariance via ALS. If, additionally, (w_k) and (v_k) follow a distribution other than the normal, then in some cases it might be possible to uniquely identify (Q, R, S) with a method that uses the third and higher sample moments of the outputs. This topic has not been considered by many researchers and is another possible avenue of future research. Kost et al. (2018) claim to be the first to give a method for estimating higher moments of the noises.

Furthermore, it bears mentioning that we have not settled the identifiability question in complete generality, as there are classes of covariance estimation methods beyond correlation and ML, such as covariance matching and Bayesian. A detailed comparison of the identifiability conditions among all classes of covariance estimation methods is another area for future research.

8.2 Tractable ALS weighting matrix calculation

Objective 2. Improve the ALS technique for covariance estimation by developing a tractable way to calculate and estimate the optimal weighting matrix for the ALS problem.

8.2.1 Summary of contributions

Of the three research objectives for the dissertation, Objective 2 is the least ambiguous in nature. Therefore, we can summarize our contributions toward this objective by noting that we accomplish exactly what the objective states. Chapter 4 presents the first generally tractable method to calculate the optimal ALS weighting matrix and Chapter 6 presents three ALS-based covariance estimation algorithms (specifically, Algorithms 2–4) which are enabled by the novel method for calculation of the optimal weighting matrix.

8.2.2 Suggestions for future work

Extend results to other correlation methods

As it is stated, Objective 2 has been completely satisfied. However, there exist least squares-based covariance estimation methods other than ALS; a notable example is

the measurement difference autocovariance method (Duník et al., 2018). Most such methods assume an identity matrix for the weighting in the least squares problem. It remains to be seen whether or not the optimal weighting matrix can be tractably calculated and estimated for such methods in a similar manner to how we have achieved this objective for ALS.

8.3 Comparison of covariance estimation algorithms

Objective 3. Compare the ALS and ML approaches to covariance estimation, both in terms of the statistical properties which are theoretically guaranteed by various algorithms and in terms of empirically observed performance of the algorithms.

8.3.1 Summary of contributions

Chapter 6 presents six ALS-based covariance estimation algorithms and provides theoretical analysis of each. Perhaps most notable are Algorithms 3 and 4, which use different approaches to estimate the optimal ALS weighting matrix from data. These algorithms would not be possible to implement without the advances achieved in Chapter 4. For Algorithms 3 and 4, we do not make any claim as to whether they produce unbiased, biased, consistent, or inconsistent covariance estimates; the algorithms are complicated enough that such theoretical analysis of these statistical properties is difficult. However, as demonstrated in Chapter 7, this difficulty does not imply that it is necessarily advisable to avoid using these algorithms.

Chapter 7 presents 17 specific covariance estimation methods, and applies each method to 100 different data sets. Twelve of the methods are based on the ALS algorithms from Chapter 6 and four of the methods are ML-based. Several of the methods are not realizable in practice, as they depend on a priori knowledge of the covariance matrices and the exact noise sequences $w_{0:T_f}$ and $v_{0:T_f}$, but are included in the analysis for the sake of comparison.

Chapter 7 also points out the fact that the computation time for ML-based methods increases rapidly with the number of data points, which is not the case for ALS-based methods. This reality is a detriment to ML estimation, as it is desirable to use large data sets to attain the best possible covariance estimates. Our interest in covariance estimation stems from collaboration with practitioners in the process industries, who often work with systems for which large data sets are available from

a data historian server. ALS-based methods are more useful than ML estimation in such situations, as they can more easily be used with large data sets, but ML estimation might be preferred in applications that do not lend themselves to large data sets.

Our opinion is that ALS with iterative weight calculation, Methods 12 and 13 (which are based on Algorithm 4), are the methods that exhibit the most desirable properties. Although there is no theoretical guarantee that these methods produce estimates which are either unbiased or consistent, empirically the estimates display relatively low bias and sample variance, only slightly larger than what is achieved with ML estimation. Furthermore, since they are ALS-based, the computation time for these methods scales favorably with the number of data points.

8.3.2 Suggestions for future work

Further comparison and evaluation of covariance estimation algorithms

The covariance estimation literature is ever evolving, so new results must frequently be compared to previous results and validated on real world data to have any promise of ever being applied to real world applications. Most papers in the covariance estimation literature include a numerical example where the method of interest is applied to a data set, the data sometimes being simulated and sometimes coming from a real plant. Zagrobelny and Rawlings (2015b) highlight the necessity of testing algorithms on real data. It is common to employ an integrated disturbance model to account for otherwise unmodeled system disturbances (Rajamani et al., 2009), but in their paper, Zagrobelny and Rawlings find that they must employ a so called double integrated disturbance model in order to obtain good results from the ALS estimation procedure on the particular industrial data set which they present. See Zagrobelny (2014, Chapter 5) for more discussion of double integrated disturbance models.

Previous research has successfully applied ALS-based methods to real plant and laboratory data. See, in addition to the aforementioned Zagrobelny and Rawlings (2015b), also Odelson et al. (2006a) and Rajamani (2007, Chapter 7), for example. Moreover, anecdotally, we have successfully solved ALS problems using randomly generated systems and data with as many as 100 states. Nevertheless, additional case studies would go a long way towards convincing industrial practitioners of the potential benefits that could be reaped in implementing ALS-based covariance estimation algorithms into their hierarchy of process control.

Although we believe that the work of Chapters 6 and 7 has resulted in some compelling conclusions, we concede that the scope of the analysis is relatively narrow. For instance, in Chapter 7, we have considered only two numerical examples: one with a two-state system and the other with a ten-state system. Furthermore, we have only used simulated data. Future work should endeavor to validate the effectiveness of the covariance estimation algorithms from this dissertation on larger systems and using real plant data. In particular, such validation should be done with Algorithm 4 from Chapter 6, ALS estimation with iterative weight calculation; as we mentioned above, our results suggest that this algorithm is the most promising covariance estimation method we consider.

In Chapter 1, we offer Duník et al. (2017) and Odelson (2003, Chapter 3) as sources which compare a wide swath of covariance estimation algorithms. We wish to call attention to one particular algorithm which we believe should be compared more thoroughly to ALS-based methods. The measurement difference autocovariance (MDA) method (Duník et al., 2018) is a correlation method which is similar to ALS in that it based on a least squares problem. MDA is conceptually similar to ALS, but the notation and bookkeeping of MDA are quite a bit more complicated than for ALS, and MDA relies on more user specified parameters than does ALS. One benefit of the MDA method is that it transforms the system outputs in a way that guarantees that the obtained covariance estimates are unbiased without appealing to Assumption 7. In contrast, Assumption 7 must hold for the ALS-based Algorithms 1 and 2 from Chapter 6 to provide unbiased estimates.

We mention near the beginning of Section 3.1 that it is often necessary to discard a bit of the data when using ALS-based methods in order to ensure that Assumption 7 holds, which is a drawback of the technique. If A (or $\mathbb{A}$, if using the $\mathscr{K}$-error system) is not sufficiently stable, one might have to discard a large amount of data to ensure that Assumption 7 holds. This scenario is undesirable, especially in applications that do not lend ready access to large data sets. MDA avoids this drawback, but at present there is not any evidence which suggests that this advantage of MDA is strong enough that MDA might be preferred over ALS in applications, especially in light of the recent advances that have been made in estimating the optimal weighting matrix for ALS espoused in Chapters 4, 6 and 7 of this dissertation and in Arnold and Rawlings (2020b). However, we do not mean to say that ALS should be preferred over MDA either—we believe that the two methods present different strengths and weaknesses relative to one another, and so a detailed comparison of these two methods in particular might be a fruitful subject for future research.

Improve understanding of practical aspects of ALS-based estimation

Our opinion is that the covariance estimation literature is relatively mature. The most important questions of the subject have been answered, in the sense that there exist a number of effective covariance estimation methods which are well known among researchers and practitioners. However, there is always room for improvements to methods and increased understanding of how to use them productively. Let us present a few questions related to the practical aspects of ALS-based estimation for which further investigation could provide practitioners with a more complete understanding of when, why, and how to use ALS in industrial settings:

- When using the $\mathcal{K}$-error system, how much does the choice of the initial predictor gain, $\mathcal{K}$, affect the results? Does $\mathcal{K}$ matter at all if the optimal weighting matrix is used for the ALS problem?
- For constrained ALS problems, if the constraints are active at the solution to the (Q, R, S) problem, are the constraints necessarily active at the solution to the (Φ, Θ) (or (Φ, Θ)) problem? And vice versa?
- In Chapter 7, we see that constrained ALS problems generally take longer to solve than their unconstrained counterparts. Therefore, it is beneficial to simply solve the unconstrained version of the ALS problem if one knows a priori that it is likely to yield a solution which satisfies the semidefinite constraints anyways. Thus, for constrained ALS problems, how often and under what circumstances can we expect the constraints to be active at the solution?
- How big of an impact can the $S = 0$ assumption have? In other words, how are the results affected if one estimates $(Q, R, 0)$ when $S \neq 0$, or, conversely, (Q, R, S) when $S = 0$?
- A common technique with least squares-based estimation techniques is to add regularization terms to the objective function of the optimization problem, especially if the unregularized problem does not have a unique solution. We do not discuss regularized ALS in this dissertation aside from a brief allusion to the topic in Section 3.4, but the Covest software package mentioned in Chapter 7 is capable of solving a wide variety of regularized ALS problems. In what ways can adding (or not adding) regularization terms to the ALS problem affect the produced estimates?

Each of these questions, considered individually, does not represent a particularly high value research question. Answering these questions, and other similar ones, collectively, however, would be a noteworthy contribution to the literature.

8.4 Future work: extension to linear time-varying and nonlinear systems

In this dissertation, our focus is exclusively linear time-invariant systems. Another area for future work is to determine if (and if so, to what extent) the results of this dissertation may be extended to linear time-varying and nonlinear systems. Covariance estimation for nonlinear systems has been considered by other authors. For example, Rajamani (2007, Chapter 7), Lima and Rawlings (2011), Lima et al. (2013), Ge and Kerrigan (2017a), and Arnold and Rawlings (2018) study ALS-based methods for these classes of systems, and Bavdekar et al. (2011), Bavdekar and Patwardhan (2012) and Ge and Kerrigan (2017b) study ML-based methods. We offer comments about the state each of the three main objectives of this dissertation as they relate to linear time-varying and nonlinear systems:

- Objective 1: Arnold and Rawlings (2018) briefly touch on covariance identifiability for linear time-varying and nonlinear systems, but a more comprehensive study would be useful.
- Objective 2: It remains to be seen whether or not the optimal weighting matrix for linear time-varying and nonlinear ALS methods can be tractably calculated and estimated in a way similar to what is achieved in this dissertation for linear time-invariant systems.
- Objective 3: As far as we are aware, there exists no comprehensive comparison of covariance estimation algorithms specifically for linear time-varying and nonlinear systems

Do not worry about your difficulties in mathematics. I can assure you mine are still greater.

Albert Einstein

APPENDIX A

UNIFIED LINEAR ESTIMATION AND LEAST SQUARES

This appendix provides background information on the linear regression model and the minimum variance affine unbiased estimator (MVAUE) which is the basis for the ALS method of covariance estimation. The appendix is organized as follows:

- Section A.1 introduces the linear regression model and describes the criteria for an estimator to be an MVAUE.
- Section A.2 provides several miscellaneous results that are necessary in the remaining sections of the appendix.
- Section A.3 presents the MVAUE for the linear regression model and discusses its connection to least squares problems.
- Section A.4 offers commentary on the literature related to this topic.

A.1 The linear regression model

Consider the linear regression model

$$y = X\beta + \varepsilon \tag{A.1}$$
$$\mathrm{E}[\varepsilon] = 0 \qquad \mathrm{E}[\varepsilon\varepsilon^{\mathrm{T}}] = V$$

in which y is the measurement vector, X is the model matrix, β is the parameter vector, and ε is the noise vector. We make no assumptions about the linear independence of the columns of either X or V, nor about the probability distribution of the random variable ε. We consider X to be a known constant and β to be an unknown constant that we would like to estimate using y.

Our goal is to find a minimum variance affine unbiased estimator (MVAUE) of β. Precisely, $\hat{\beta}$ is an MVAUE of β if it has the following properties:

- *Affine:* The estimator takes the form of an affine function of y, i.e., $\hat{\beta} = My + c$.
- *Unbiased:* The estimator satisfies $\mathrm{E}[\hat{\beta}] = \beta$ for all β. For an affine estimator $\hat{\beta} = My + c$, $\mathrm{E}[\hat{\beta}] = M\,\mathrm{E}[y] + c = MX\beta + c$, so unbiasedness requires $MX = I$ and $c = 0$.
- *Minimum variance:* The estimator satisfies $\mathrm{cov}(\hat{\beta}^*) - \mathrm{cov}(\hat{\beta}) \geq 0$ for any affine unbiased estimator $\hat{\beta}^*$.

Note that $\hat{\beta}$ is generally a function of y and thus might be more appropriately denoted $\hat{\beta}(y)$. For the sake of presentation, we use the former notation.

A.2 Miscellaneous results

Next, we provide some results that are needed in the remainder of the appendix.

Theorem A.1 (Magnus and Neudecker (2019), Theorem 11.35)**.** *Suppose W is a positive semidefinite matrix. Then for arbitrary q,*

$$\beta^* = (X^{\mathrm{T}}WX)^{+}X^{\mathrm{T}}Wy + [I - (X^{\mathrm{T}}WX)^{+}X^{\mathrm{T}}WX]q$$

is a global minimizer of the linear least squares function

$$f(\beta) \coloneqq \|X\beta - y\|_W^2$$

Lemma A.1. *For any matrix A,* $\mathrm{col}\, A = \mathrm{col}\, AA^{\mathrm{T}}$.

Proof. It is obvious that $\mathrm{col}\, AA^{\mathrm{T}} \subseteq \mathrm{col}\, A$. The converse $\mathrm{col}\, A \subseteq \mathrm{col}\, AA^{\mathrm{T}}$ follows from the identity $A = AA^{\mathrm{T}}(A^{+})^{\mathrm{T}}$ (Magnus and Neudecker, 2019, Theorem 2.5). ■

Lemma A.2. *Let A and B be positive semidefinite matrices of the same size. Then* $\mathrm{col}\, A \subseteq \mathrm{col}(A + B)$.

Proof. Let $x \in \operatorname{null}(A+B)$. Then

$$(A+B)x = 0 \implies x^{\mathrm{T}}Ax + x^{\mathrm{T}}Bx = 0 \implies x^{\mathrm{T}}Ax = 0 \implies Ax = 0$$

The final implication in the above follows because A is positive semidefinite (Horn and Johnson, 2013, Observation 7.1.6). We have shown

$$\operatorname{null}(A+B) = (\operatorname{col}(A+B))^{\perp} \subseteq (\operatorname{col} A)^{\perp} = \operatorname{null} A$$

from which the result follows. ■

Lemma A.3. *Suppose matrix X has LI columns, matrix W is positive semidefinite, and $\operatorname{col} W \supseteq \operatorname{col} X$. Then $X^{\mathrm{T}}WX$ is positive definite.*

Proof. Write the singular value decomposition of W:

$$W = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} U_1^{\mathrm{T}} \\ U_2^{\mathrm{T}} \end{bmatrix} = U_1 \Sigma U_1^{\mathrm{T}}$$

Then

$$\operatorname{col} U_1 = \operatorname{col} W \supseteq \operatorname{col} X \implies X = U_1 Y \text{ for some matrix } Y$$

Furthermore, Y has LI columns. If Y had LD columns, then $X = U_1 Y$ would imply that X has LD columns, a contradiction. Finally,

$$X^{\mathrm{T}}WX = Y^{\mathrm{T}}U_1^{\mathrm{T}}U_1 \Sigma U_1^{\mathrm{T}}U_1 Y = Y^{\mathrm{T}}\Sigma Y > 0$$ ■

A.3 The minimum variance affine unbiased estimator

The next theorem gives the main result of interest:

Theorem A.2 (Magnus and Neudecker (2019), Theorem 13.13)**.** *Suppose y is generated by the linear model* (A.1)*. The MVAUE of $Z\beta$ (a linear transformation of β) exists if and only if $\operatorname{col} Z^{\mathrm{T}} \subseteq \operatorname{col} X^{\mathrm{T}}$. In this case, the MVAUE is unique and given by*

$$\widehat{Z\beta} = Z(X^{\mathrm{T}}V_0^{+}X)^{+}X^{\mathrm{T}}V_0^{+}y$$
$$\operatorname{cov}(\widehat{Z\beta}) = Z[(X^{\mathrm{T}}V_0^{+}X)^{+} - I]Z^{\mathrm{T}}$$

where $V_0 := V + XX^{\mathrm{T}}$. Moreover, if $\operatorname{col} X \subseteq \operatorname{col} V$, then the estimator simplifies to

$$\widehat{Z\beta} = Z(X^{\mathrm{T}}V^{+}X)^{+}X^{\mathrm{T}}V^{+}y$$
$$\operatorname{cov}(\widehat{Z\beta}) = Z(X^{\mathrm{T}}V^{+}X)^{+}Z^{\mathrm{T}}$$

Corollary A.1. *An MVAUE of β exists if and only if X has LI columns.*

Proof. In this case, $Z = I$. The existence condition is $\operatorname{col} I \subseteq \operatorname{col} X^{\mathrm{T}}$, which is equivalent to X having LI columns. ■

The MVAUE has a strong connection to linear least squares problems which is elucidated by the following corollary:

Corollary A.2. *Take $Z = I$ and assume X has LI columns. Then the MVAUE $\hat{\beta}$ is the unique solution to the following linear least squares problem:*

$$\hat{\beta} = \arg\min_{\beta} \|X\beta - y\|_{V_0^+}^2$$

If $\operatorname{col} X \subseteq \operatorname{col} V$, this problem can be simplified to

$$\hat{\beta} = \arg\min_{\beta} \|X\beta - y\|_{V^+}^2$$

Proof. That $\hat{\beta}$ minimizes the linear least squares function is immediately apparent by comparing Theorems A.1 and A.2. This minimizer is unique if and only if the Hessian $X^{\mathrm{T}}V_0^+X$ (or, in the second case, $X^{\mathrm{T}}V^+X$) is invertible. In the first case, using Lemmas A.1 and A.2, we have $\operatorname{col} V_0^+ = \operatorname{col} V_0 \supseteq \operatorname{col} XX^{\mathrm{T}} = \operatorname{col} X$. Therefore, applying Lemma A.3 gives $X^{\mathrm{T}}V_0^+X > 0$. Similarly, in the second case, we have $\operatorname{col} V = \operatorname{col} V^+ \supseteq \operatorname{col} X$, so Lemma A.3 implies $X^{\mathrm{T}}V^+X > 0$. ■

It is known that the choice of V_0 in Theorem A.2 is not unique:

Theorem A.3. *Define $V_1 := V + XEX^{\mathrm{T}}$. Then $(X^{\mathrm{T}}V_1^+X)^+X^{\mathrm{T}}V_1^+$ does not depend on E, provided that $E > 0$.*[1]

Proof. We begin by writing the singular value decomposition of X:

$$X = USW^{\mathrm{T}} = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} W_1^{\mathrm{T}} \\ W_2^{\mathrm{T}} \end{bmatrix} = U_1 \Sigma W_1^{\mathrm{T}}$$

[1]The author would like to thank Jim Rawlings of UCSB for help in attaining the proof which is shown for Theorem A.3. Most of the insight which led to the method of proof which is shown here is owed to him. Furthermore, Professor Rawlings has published his work on the topic as extra exercises to his textbook, Graham and Rawlings (2013), which are available at `sites.engineering.ucsb.edu/~jbraw/principles/`.

Then

$$\begin{aligned} V_1 &= V + XEX^{\mathrm{T}} \\ &= UU^{\mathrm{T}}VUU^{\mathrm{T}} + USW^{\mathrm{T}}EWSU^{\mathrm{T}} \\ &= U(U^{\mathrm{T}}VU + SW^{\mathrm{T}}EWS)U^{\mathrm{T}} \\ &= U \begin{bmatrix} \Sigma W_1^{\mathrm{T}} E W_1 \Sigma + U_1^{\mathrm{T}} V U_1 & U_1^{\mathrm{T}} V U_2 \\ U_2^{\mathrm{T}} V U_1 & U_2^{\mathrm{T}} V U_2 \end{bmatrix} U^{\mathrm{T}} \end{aligned}$$

Using the fact that $(ABC)^+ = C^{\mathrm{T}}B^+A^{\mathrm{T}}$ if A and C are orthogonal (a fact which may be verified easily), we get

$$V_1^+ = U \underbrace{\begin{bmatrix} \Sigma W_1^{\mathrm{T}} E W_1 \Sigma + U_1^{\mathrm{T}} V U_1 & U_1^{\mathrm{T}} V U_2 \\ U_2^{\mathrm{T}} V U_1 & U_2^{\mathrm{T}} V U_2 \end{bmatrix}^+}_{=: \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}} U^{\mathrm{T}}$$

Our next task is to find expressions for A_{11}, A_{12}, A_{21}, and A_{22}. To this end, define

$$B := \Sigma W_1^{\mathrm{T}} E W_1 \Sigma + B_1 \tag{A.2}$$

$$\begin{aligned} B_1 &:= U_1^{\mathrm{T}} V U_1 - U_1^{\mathrm{T}} V U_2 (U_2^{\mathrm{T}} V U_2)^+ U_2^{\mathrm{T}} V U_1 \\ &= \begin{bmatrix} I & -U_1^{\mathrm{T}} V U_2 (U_2^{\mathrm{T}} V U_2)^+ \end{bmatrix} \begin{bmatrix} U_1^{\mathrm{T}} V U_1 & U_1^{\mathrm{T}} V U_2 \\ U_2^{\mathrm{T}} V U_1 & U_2^{\mathrm{T}} V U_2 \end{bmatrix} \begin{bmatrix} I \\ -(U_2^{\mathrm{T}} V U_2)^+ U_2^{\mathrm{T}} V U_1 \end{bmatrix} >= 0 \end{aligned}$$

The matrix B_1 can be considered a generalized Schur complement. Next, we have $\Sigma W_1^{\mathrm{T}} E W_1 \Sigma > 0$, since $E > 0$ and $W_1\Sigma$ has LI columns. Therefore $B > 0$, and hence B is invertible. Thus, the expressions for A_{11}, A_{12}, A_{21}, and A_{22} are

$$\begin{aligned} A_{11} &= B^{-1} \\ A_{12} &= -A_{11} U_1^{\mathrm{T}} V U_2 (U_2^{\mathrm{T}} V U_2)^+ \\ A_{21} &= A_{12}^{\mathrm{T}} \\ A_{22} &= (U_2^{\mathrm{T}} V U_2)^+ + (U_2^{\mathrm{T}} V U_2)^+ U_2^{\mathrm{T}} V U_1 A_{11} U_1^{\mathrm{T}} V U_2 (U_2^{\mathrm{T}} V U_2)^+ \end{aligned}$$

It can be verified directly that these values of A_{11}, A_{12}, A_{21}, and A_{22} satisfy the four Moore-Penrose inverse conditions. We leave the details as an exercise for the reader, but offer the following hints: use Corollary 2.10 and the identity

$$U_1^{\mathrm{T}} V U_2 (U_2^{\mathrm{T}} V U_2)^+ U_2^{\mathrm{T}} V U_1 = \Sigma W_1^{\mathrm{T}} E W_1 \Sigma + U_1^{\mathrm{T}} V U_1 - A_{11}^{-1}$$

For future reference, observe that $A_{11}^{-1}A_{12} = -U_1^{\mathrm{T}}VU_2(U_2^{\mathrm{T}}VU_2)^{+}$. We now have all we need to attain the final result. Observe

$$\begin{aligned} X^{\mathrm{T}}V_1^{+}X &= WSU^{\mathrm{T}}U\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22}\end{bmatrix}U^{\mathrm{T}}USW^{\mathrm{T}} \\ &= WS\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22}\end{bmatrix}SW^{\mathrm{T}} \\ &= W\begin{bmatrix} \Sigma A_{11}\Sigma & 0 \\ 0 & 0\end{bmatrix}W^{\mathrm{T}} \end{aligned}$$

Then

$$(X^{\mathrm{T}}V_1^{+}X)^{+} = W\begin{bmatrix} \Sigma A_{11}\Sigma & 0 \\ 0 & 0\end{bmatrix}^{+}W^{\mathrm{T}} = W\begin{bmatrix} \Sigma^{-1}A_{11}^{-1}\Sigma^{-1} & 0 \\ 0 & 0\end{bmatrix}W^{\mathrm{T}}$$

Finally,

$$\begin{aligned} (X^{\mathrm{T}}V_1^{+}X)^{+}X^{\mathrm{T}}V_1^{+} &= W\begin{bmatrix} \Sigma^{-1}A_{11}^{-1}\Sigma^{-1} & 0 \\ 0 & 0\end{bmatrix}W^{\mathrm{T}}WSU^{\mathrm{T}}U\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22}\end{bmatrix}U^{\mathrm{T}} \\ &= \begin{bmatrix} W_1 & W_2\end{bmatrix}\begin{bmatrix} \Sigma^{-1}A_{11}^{-1}\Sigma^{-1} & 0 \\ 0 & 0\end{bmatrix}\begin{bmatrix} \Sigma & 0 \\ 0 & 0\end{bmatrix}\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22}\end{bmatrix}\begin{bmatrix} U_1^{\mathrm{T}} \\ U_2^{\mathrm{T}}\end{bmatrix} \\ &= \begin{bmatrix} W_1 & W_2\end{bmatrix}\begin{bmatrix} \Sigma^{-1} & \Sigma^{-1}A_{11}^{-1}A_{12} \\ 0 & 0\end{bmatrix}\begin{bmatrix} U_1^{\mathrm{T}} \\ U_2^{\mathrm{T}}\end{bmatrix} \\ &= \begin{bmatrix} W_1 & W_2\end{bmatrix}\begin{bmatrix} \Sigma^{-1} & -\Sigma^{-1}U_1^{\mathrm{T}}VU_2(U_2^{\mathrm{T}}VU_2)^{+} \\ 0 & 0\end{bmatrix}\begin{bmatrix} U_1^{\mathrm{T}} \\ U_2^{\mathrm{T}}\end{bmatrix} \\ &= W_1\Sigma^{-1}U_1^{\mathrm{T}} - W_1\Sigma^{-1}U_1^{\mathrm{T}}VU_2(U_2^{\mathrm{T}}VU_2)^{+}U_2^{\mathrm{T}} \\ &= X^{+} - X^{+}VU_2(U_2^{\mathrm{T}}VU_2)^{+}U_2^{\mathrm{T}} \end{aligned}$$

We see that E does not appear in the final expression for $(X^{\mathrm{T}}V_1^{+}X)^{+}X^{\mathrm{T}}V_1^{+}$. ■

Note that $V_0 = V_1$ if $E = I$.

A.4 Comments on the literature

Theorem A.2 with the additional assumption that V is invertible is due to Aitken (1935) (before Aitken's paper, the only case that was known was that where V is a scalar multiple of the identity matrix). The case shown in Theorem A.2, where V may be singular, is due to C. Radhakrishna Rao (he of the famed Cramér-Rao

bound). Rao published a series of papers on the subject, the most notable perhaps being Rao (1971). The book Rao (1973a, Chapter 4, Section i in particular) gives a full derivation of the results. Rao's derivations of the results include the generalization with E that is described in Theorem A.3.

Rao's works, while comprehensive, are terse in places and can be difficult to follow without sufficient background knowledge of the subject. Magnus and Neudecker (2019, Chapter 13) give a more expository presentation of the MVAUE. It is worth noting that these two sets of authors draw on different motivating principles in their derivations. Rao is guided by the connection between the MVAUE and least squares seen in earlier works (such as Aitken's), and he seeks an estimator of this form from the beginning. Magnus and Neudecker on the other hand begin by insisting only that the estimator be affine and unbiased, and proceed by minimizing the trace of the variance among that class of estimators. The connection between the obtained MVAUE and the least squares problem is noted after the fact.

We remark that Rao's works use a broader notion of matrix inverse than the Moore-Penrose inverse, namely, the so called generalized inverse (or g-inverse). A matrix X^- is said to be a g-inverse of X if it satisfies $XX^-X = X$ (Rao, 1973a, Section 1b.5). The Moore-Penrose inverse clearly has this property, and it is used here because it is the most practical unique g-inverse in most cases. For instance, the Moore-Penrose inverse is unique for all matrices, while the g-inverse is unique only for square invertible matrices. Rao and Mitra (1971) provide a comprehensive review of a wide variety of generalized matrix inverses.

Moreover, Rao's works indicate that we can weaken the assumption on E in Theorem A.3 and achieve the same result:

Conjecture A.1. *Theorem A.3 holds if the assumption on E is weakened from $E > 0$ to $E \geq 0$ and $\operatorname{col} X \subseteq \operatorname{col}(V + XEX^{\mathrm{T}})$.*

In addition to the aforementioned works by Rao, Rao (1973b) is also of interest on this matter. Magnus and Neudecker (2019) also mention the result on page 315, but they do not provide a proof. We found Rao's proofs on this topic confusing and unconvincing, which was our motivation for seeking our own proof of Theorem A.3. However, we have not been able to achieve a proof of Conjecture A.1. Fortunately, the $E > 0$ case is all that is needed in the main body of this dissertation (see Theorems 3.3 and 3.4).

Let us describe the difficulty we encountered in attempting to prove Conjecture A.1 along similar lines as we used in proving Theorem A.3. With $E \geq 0$, we can only assume $\Sigma W_1^{\mathrm{T}} E W_1 \Sigma \geq 0$ rather than > 0. Therefore, we cannot immediately

conclude that the matrix B in (A.2) is invertible. We suspect that B does remain invertible with the weaker conditions on E, as we have not found a counterexample to the contrary. If indeed B remains invertible, then the proof of Conjecture A.1 goes through identically to the remainder of the proof for Theorem A.3. The following facts may be helpful to anyone who wishes to attempt to prove that B is invertible:

1. Observe:

$$\begin{aligned}\Sigma W_1^{\mathrm{T}} E W_1 \Sigma + U_1^{\mathrm{T}} V U_1 &= U_1^{\mathrm{T}} U_1 \Sigma W_1^{\mathrm{T}} E W_1 \Sigma U_1^{\mathrm{T}} U_1 + U_1^{\mathrm{T}} V U_1 \\ &= U_1^{\mathrm{T}} (V + X E X^{\mathrm{T}}) U_1\end{aligned}$$

2. If $X \neq 0$, then $\operatorname{col} X \subseteq \operatorname{col}(V + XEX^{\mathrm{T}}) \implies U_1^{\mathrm{T}}(V + XEX^{\mathrm{T}})U_1 > 0$, since $\operatorname{col} U_1 = \operatorname{col} X$ (see Lemma A.3). Because $\operatorname{col}(V+XEX^{\mathrm{T}}) = \operatorname{col}(V+XEX^{\mathrm{T}})^{+}$, it follows that $U_1^{\mathrm{T}}(V + XEX^{\mathrm{T}})^{+}U_1 > 0$ with the same reasoning.

3. It follows from the results in Carlson et al. (1974) that B is invertible if and only if the matrices

$$\begin{bmatrix} U_1^{\mathrm{T}}(V + XEX^{\mathrm{T}})U_1 & U_1^{\mathrm{T}} V U_2 \\ U_2^{\mathrm{T}} V U_1 & U_2^{\mathrm{T}} V U_2 \end{bmatrix} \quad \text{and} \quad U_2^{\mathrm{T}} V U_2$$

have the same number of zero eigenvalues. One way to establish this result would be by proving the following conjecture:

Conjecture A.2. *The following result holds:*

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \operatorname{null} \begin{bmatrix} U_1^{\mathrm{T}}(V + XEX^{\mathrm{T}})U_1 & U_1^{\mathrm{T}} V U_2 \\ U_2^{\mathrm{T}} V U_1 & U_2^{\mathrm{T}} V U_2 \end{bmatrix}$$
$$\iff x_1 = 0 \ \textit{and} \ x_2 \in \operatorname{null}(U_2^{\mathrm{T}} V U_2)$$

The $\impliedby$ part of Conjecture A.2 is easy to establish, but we have been unsuccessful in our attempts to prove or disprove the $\implies$ direction.

To summarize, establishing Conjecture A.2 would be sufficient to also establish Conjecture A.1.

BIBLIOGRAPHY

A. C. Aitken. On least squares and linear combinations of observations. *Proceedings of the Royal Society of Edinburgh*, 55:42–48, 1935.

Bernt M. Åkesson, John Bagterp Jørgensen, Niels Kjølstad Poulsen, and Sten Bay Jørgensen. A generalized autocovariance least-squares method for Kalman filter tuning. *Journal of Process Control*, 18:769–779, 2008.

Brian D. O. Anderson and John B. Moore. *Optimal Filtering*. Prentice-Hall, Englewood Cliffs, N. J., 1979.

Travis J. Arnold and James B. Rawlings. Uniqueness conditions for ALS problems. 6th IFAC Conference on Nonlinear Model Predictive Control NMPC 2018, Madison, WI, August 2018.

Travis J. Arnold and James B. Rawlings. On the identifiability of noise covariance matrices for time-invariant linear systems, 2020a. Submitted to *IEEE Transactions on Automatic Control*.

Travis J. Arnold and James B. Rawlings. Tractable calculation and estimation of the optimal weighting matrix for ALS problems, 2020b. Submitted to *IEEE Transactions on Automatic Control*.

V. A. Bavdekar, A. P. Deshpande, and S. C. Patwardhan. Identification of process and measurement noise covariance for state and parameter estimation using extended Kalman filter. *Journal of Process Control*, 21:585–601, 2011.

Vinay A. Bavdekar and Sachin C. Patwardhan. Development of grey box state estimators for systems subjected to time correlated unmeasured disturbances. *Journal of Process Control*, 22(9):1543–1558, 2012.

Sean Borman. The expectation maximization algorithm—a short tutorial, July 2004. Available: seanborman.com/publications/EM_algorithm.pdf. Updated January 2009.

S. P. Boyd and L Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

David Carlson, Emile Haynsworth, and Thomas Markham. A generalization of the Schur complement by means of the Moore-Penrose inverse. *SIAM Journal on Applied Mathematics*, 26(1):169–175, 1974.

A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B*, 39:1–38, 1977.

Chuong B. Do and Serafim Batzoglou. What is the expectation maximization algorithm? *Nature Biotechnology*, 26(8):897–899, 2008.

Jindřich Duník, Ondřej Straka, Oliver Kost, and Jindřich Havlík. Noise covariance matrices in state-space models: A survey and comparison of estimation methods—part I. *International Journal of Adaptive Control and Signal Processing*, 31(11):1505–1543, 2017.

Jindřich Duník, Oliver Kost, and Ondřej Straka. Design of measurement difference autocovariance method for estimation of process and measurement noise covariances. *Automatica*, 90:16–24, 2018.

Pierre L. Faurre. Stochastic realization algorithms. In Raman K. Mehra and Dimitri G. Lainiotis, editors, *System Identification Advances and Case Studies*, volume 126 of *Mathematics in Science and Engineering*, pages 1–25. Elsevier, 1976.

Ming Ge and Eric C. Kerrigan. Noise covariance identification for time-varying and nonlinear systems. *International Journal of Control*, 90(9):1903–1915, 2017a.

Ming Ge and Eric C. Kerrigan. Noise covariance identification for nonlinear systems using expectation maximization and moving horizon estimation. *Automatica*, 77: 336–343, 2017b.

Zoubin Ghahramani and Geoffrey E. Hinton. Parameter estimation for linear dynamical systems. Technical Report CRG-TR-96-2, Department of Computer Science, University of Toronto, 1996.

Ghazal A. Ghazal and Heinz Neudecker. On second-order and fourth-order moments of jointly distributed random matrices: a survey. *Linear Algebra and its Applications*, 321:61–93, 2000.

Stuart Gibson and Brett Ninness. Robust maximum-likelihood estimation of multivariable dynamic systems. *Automatica*, 41(10):1667–1682, 2005.

Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, MD, fourth edition, 2013.

Michael D. Graham and James B. Rawlings. *Modeling and Analysis Principles for Chemical and Biological Engineers*. Nob Hill Publishing, Madison, WI, 2013.

John A. Gubner. *Probability and Random Processes for Electrical and Computer Engineers*. Cambridge University Press, New York, 2006.

M. L. J. Hautus. Controllability and stabilizability of sampled systems. *IEEE Transactions on Automatic Control*, 17(4):528–531, August 1972.

Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, second edition, 2013.

Tohru Katayama. *Subspace Methods for System Identification*. Springer-Verlag, London, 2005.

Steven M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*, volume 1. Prentice Hall, Upper Saddle River, NJ, 1993.

Oliver Kost, Jindřich Duník, and Ondřej Straka. Noise moment and parameter estimation of state-space model. In *18th IFAC Symposium on System Identification*, pages 891–896, Stockholm, Sweden, July 2018.

Sandip K. Lahiri. *Multivariable Predictive Control: Applications in Industry*. John Wiley & Sons, Inc., Hoboken, NJ, 2017.

Weichang Li and Thomas A. Badgwell. Structured covariance estimation for state prediction. In *53rd IEEE Conference on Decision and Control*, pages 296–303, Los Angeles, CA, December 15–17 2014.

Fernando V. Lima and James B. Rawlings. Nonlinear stochastic modeling to improve state estimation in process monitoring and control. *AIChE Journal*, 57(4):996–1007, 2011.

Fernando V. Lima, Murali R. Rajamani, Tyler A. Soderstrom, and James B. Rawlings. Covariance and state estimation of weakly observable systems: Application to polymerization processes. *IEEE Transactions on Control Systems Technology*, 21(4): 1249–1257, July 2013.

Lennart Ljung. *System Identification: Theory for the User*. Prentice Hall, New Jersey, second edition, 1999.

Jan R. Magnus and H. Neudecker. The commutation matrix: Some properties and applications. *Annals of Statistics*, 7(2):381–394, March 1979.

Jan R. Magnus and Heinz Neudecker. *Matrix Differential Calculus with Applications in Statistics and Econometrics*. John Wiley & Sons Ltd., Hoboken, NJ, third edition, 2019.

Geoffrey J. McLachlan and Thriyambakam Krishnan. *The EM Algorithm and Extensions*. John Wiley & Sons, Inc., Hoboken, NJ, second edition, 2008.

R. K. Mehra. On the identification of variances and adaptive Kalman filtering. *IEEE Transactions on Automatic Control*, 15(12):175–184, 1970.

Whitney K. Newey and Daniel L. McFadden. Large sample estimation and hypothesis testing. In Robert F. Engle and Daniel L. McFadden, editors, *Handbook of Econometrics*, volume 4, chapter 36, pages 2111–2245. Elsevier North Holland, Amsterdam, 1994.

Brett Ninness and Stuart Gibson. The EM algorithm for multivariable dynamic system estimation. Technical Report EE200101, Department of Electrical and Computer Engineering, University of Newcastle, Australia, 2001.

Brian J. Odelson. *Estimating Disturbance Covariances From Data for Improved Control Performance*. PhD thesis, University of Wisconsin–Madison, 2003.

Brian J. Odelson, Alexander Lutz, and James B. Rawlings. The autocovariance least-squares method for estimating covariances: Application to model-based control of chemical reactors. *IEEE Control Systems Technology*, 14(3):532–541, May 2006a.

Brian J. Odelson, Murali R. Rajamani, and James B. Rawlings. A new autocovariance least-squares method for estimating noise covariances. *Automatica*, 42(2):303–308, February 2006b.

Babatunde A. Ogunnaike and W. Harmon Ray. *Process Dynamics, Modeling, and Control*. Oxford University Press, New York, 1994.

S. J. Qin. An overview of subspace identification. *Computers & Chemical Engineering*, 30:1502–1513, 2006.

S. Joe Qin and Thomas A. Badgwell. A survey of industrial model predictive control technology. *Control Engineering Practice*, 11(7):733–764, 2003.

Murali R. Rajamani and James B. Rawlings. Estimation of the disturbance structure from data using semidefinite programming and optimal weighting. *Automatica*, 45(1):142–148, 2009.

Murali R. Rajamani, James B. Rawlings, and S. Joe Qin. Achieving state estimation equivalence for misassigned disturbances in offset-free model predictive control. *AIChE Journal*, 55(2):396–407, February 2009.

Murali Rajaa Rajamani. *Data-based Techniques to Improve State Estimation in Model Predictive Control*. PhD thesis, University of Wisconsin–Madison, October 2007.

C. Radhakrishna Rao. Unified theory of linear estimation. *Sankhyā: The Indian Journal of Statistics, Series A*, 33(4):371–394, 1971.

C. Radhakrishna Rao. *Linear Statistical Inference and Its Applications*. John Wiley and Sons, New York, second edition, 1973a.

C. Radhakrishna Rao. Unified theory of least squares. *Communications in Statistics*, 1(1):1–8, 1973b.

C. Radhakrishna Rao and Sujit Kumar Mitra. *Generalized Inverse of Matrices and its Applications*. John Wiley & Sons, Inc., New York, 1971.

James B. Rawlings, David Q. Mayne, and Moritz M. Diehl. *Model Predictive Control: Theory, Design, and Computation*. Nob Hill Publishing, Madison, WI, 2nd edition, 2017.

Walter Rudin. *Principles of Mathematical Analysis*. McGraw-Hill, New York, third edition, 1976.

S. R. Searle. *Linear Models*. John Wiley & Sons, Inc., 1971.

Dale E. Seborg, Thomas F. Edgar, Duncan A. Mellichamp, and Francis J. Doyle. *Process Dynamics and Control*. John Wiley and Sons, New York, fourth edition, 2017.

Robert H. Shumway and David S. Stoffer. Time series smoothing and forecasting using the EM algorithm. Technical Report 27, Division of Statistics, University of California, Davis, 1981.

Robert H. Shumway and David S. Stoffer. An approach to time series smoothing and forecasting using the EM algorithm. *Journal of Time Series Analysis*, 3:253–264, 1982.

Robert H. Shumway and David S. Stoffer. *Time Series Analysis and its Applications*. Springer International Publishing, Cham, Switzerland, fourth edition, 2017.

Leonard M. Silverman. Discrete Riccati equations: Alternative algorithms, asymptotic properties, and system theory interpretations. In C. T. Leondes, editor, *Control and Dynamic Systems*, volume 12 of *Control and Dynamic Systems*, pages 313–386. Academic Press, 1976.

G. W. Stewart. On the continuity of the generalized inverse. *SIAM Journal on Applied Mathematics*, 17(1):33–45, 1969.

Edison Tse and John J. Anton. On the identifiability of parameters. *IEEE Transactions on Automatic Control*, 17(5):637–646, 1972.

A. W. van der Vaart. *Asymptotic Statistics*. Cambridge University Press, 1998.

Peter van Overschee and Bart De Moor. *Subspace Identification for Linear Systems: Theory, Implementation, Applications*. Academic Publishers, 1996.

Lieven Vandenberghe and Stephen Boyd. Semidefinite programming. *SIAM Review*, 38(1):49–95, March 1996.

C. F. Jeff Wu. On the convergence properties of the EM algorithm. *Annals of Statistics*, 11(1):95–103, 1983.

Byron M. Yu, Krishna V. Shenoy, and Maneesh Sahani. Derivation of Kalman filtering and smoothing equations. Technical report, October 2004. Available: `users.ece.cmu.edu/~byronyu/papers/derive_ks.pdf`.

Megan A. Zagrobelny and James B. Rawlings. Identification of disturbance covariances using maximum likelihood estimation. Technical Report 2014–02, TWCCC, December 2014.

Megan A. Zagrobelny and James B. Rawlings. Identifying the uncertainty structure using maximum likelihood estimation. In *American Control Conference*, pages 422–427, Chicago, IL, July 1–3, 2015a.

Megan A. Zagrobelny and James B. Rawlings. Practical improvements to autocovariance least-squares. *AIChE Journal*, 61:1840–1855, 2015b.

Megan A. Zagrobelny, Luo Ji, and James B. Rawlings. Quis custodiet ipsos custodes? *Annual Reviews in Control*, 37:260–270, 2013.

Megan Ann Zagrobelny. *MPC Performance Monitoring and Disturbance Model Identification*. PhD thesis, University of Wisconsin–Madison, December 2014.