11-3-2016

A Study of the Duality between Kalman Filters and LQR Problems

Dong-Hwan Lee
Donghwan Lee, lee1923@purdue.edu

Jianghai Hu
Purdue University, jianghai@purdue.edu

A Study of the Duality between Kalman Filters and LQR Problems

Donghwan Lee and Jianghai Hu

Abstract

The goal of this paper is to study a connection between the finite-horizon Kalman filtering and the LQR problems for discrete-time LTI systems. Motivated from the recent duality results on the LQR problem, a Lagrangian dual relation is used to prove that the Kalman filtering problem is a Lagrange dual problem of the LQR problem.

I. Introduction

In this paper, we will consider the Kalman filtering and LQR problems [1], [2], each of which is one of the most fundamental agendas in systems and control theory. There is a well-known duality between them: the Kalman filter design for a stochastic LTI system is equivalent to the LQR design problem for its dual system. The goal of this paper is to study a duality relation between them in terms of the Lagrangian duality in optimization theories [3]. There are several duality relations in systems and control theory, which have attracted much attention during the last decades. For instance, a new proof of Lyapunov's matrix inequality was developed in [4] based on the standard semidefinite programming (SDP) [5] duality. A SDP formulation of the LQR problem was presented in [6] and [7] using the SDP duality. Comprehensive studies on the SDP duality in systems and control theory, such as the Kalman-Yakubovich-Popov (KYP) lemma, the LQR problem, and the H_∞-norm computation, were provided in [8]. More recent results include the state-feedback solution to the LQR problem [9] and the generalized KYP lemma [10] derived using the Lagrangian duality.

The results of this paper are mainly motivated from the ideas in [9]. First of all, the finite-horizon LQR and Kalman filtering problems are reformulated as optimizations subject to matrix

D. Lee and J. Hu are with the Department of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47906, USA (e-mail: lee1923@purdue.edu, jianghai@purdue.edu).

equalities, which represent the covariance updates of the stochastic systems. Using the Lagrangian duality, it is proved that one problem can be converted into the other problem. It is expected that the proposed analysis can shed an insight into understanding the relations between the LQR and Kalman filtering problems. In addition, it is proved that the Riccati equation and its solution for the finite-horizon LQR problem corresponds to its dual problem and the Lagrange multipliers, respectively.

II. Preliminaries and Problem Formulation

A. Notation

The adopted notation is as follows: $\mathbb{N}$: set of nonnegative integers; $\mathbb{R}$: set of real numbers; $\mathbb{R}^n$: n-dimensional Euclidean space; $\mathbb{R}^{n\times m}$: set of all $n\times m$ real matrices; A^T: transpose of matrix A; $A \succ 0$ ($A \prec 0$, $A \succeq 0$, and $A \preceq 0$, respectively): symmetric positive definite (negative definite, positive semi-definite, and negative semi-definite, respectively) matrix A; I: identity matrix of appropriate dimensions; $\mathbb{S}^n$: symmetric $n \times n$ matrices; $\mathbb{S}^n_+$: cone of symmetric $n \times n$ positive semi-definite matrices; $\mathbb{S}^n_{++}$: symmetric $n \times n$ positive definite matrices; $\mathrm{vec}(A)$: vectorization for matrix A; $A \otimes B$: Kronecker's product of matrices A and B; $\mathrm{Tr}(A)$: trace of matrix A; $\mathrm{E}(\cdot)$: expectation operator.

B. Problem formulation

Consider the discrete-time stochastic LTI system

$$x(k+1) = Ax(k) + w(k), \quad y(k) = Cx(k) + v(k) \tag{1}$$

where $k \in \mathbb{N}$, $x(k) \in \mathbb{R}^n$ is the state vector, $y(k) \in \mathbb{R}^m$ is the output vector, $w(k) \in \mathbb{R}^n$ and $v(k) \in \mathbb{R}^m$ are independent Gaussian random vectors with zero mean and covariance matrices $Q \in \mathbb{S}^n_+$ and $R \in \mathbb{S}^n_{++}$, respectively, i.e., $w(k) \sim \mathcal{N}(0, Q)$ and $v(k) \sim \mathcal{N}(0, R)$. The initial state $x(0) \in \mathbb{R}^n$ is also an independent Gaussian random vector with zero mean and covariance $Q_f \in \mathbb{S}^n_+$. Consider the Kalman filter

$$\hat{x}(k+1) = A\hat{x}(k) + L_k(C\hat{x}(k) - y(k)) \tag{2}$$

where $\hat{x}(k) \in \mathbb{R}^n$ is the state estimation, L_k is the Kalman gain over the finite-horizon $k \in \{0, 1, \ldots, N-1\}$ given by

$$L_k = -AP_kC^T(R + CP_kC^T)^{-1}. \tag{3}$$

and $\{P_k\}_{k=0}^N$ is a solution to the Riccati equation

$$\begin{aligned} P_{k+1} &= AP_kA^T + Q - AP_kC^T(R + CP_kC^T)^{-1}CP_kA^T \\ &= (A + L_kC)P_k(A + L_kC)^T + Q + L_kRL_k^T, \quad P_0 = Q_f. \end{aligned} \tag{4}$$

Each P_k can be viewed as the covariance matrix of the estimation error at time k defined by $e(k) := \hat{x}(k) - x(k)$. The estimation error system is given by

$$e(k+1) = (A + L_kC)e(k) - w(k) - L_kv(k),$$

where $e(0) \in \mathbb{R}^n$ is a Gaussian random vector $e(0) \sim \mathcal{N}(0,\, Q_f)$. It can be represented by

$$e(k+1) = (A + L_kC)e(k) + \phi(k), \tag{5}$$

where $\phi(k) = -w(k) - L_kv(k)$ is Gaussian random vector $\phi(k) \sim \mathcal{N}(0,\, Q + L_kRL_k^T)$. From the duality between the Kalman filtering problem and the LQR problem, the equations (2)-(4) are equivalent to the Riccati equation for the LQR problem of the dual system

$$\xi(k+1) = A^T\xi(k) + C^Tu(k), \tag{6}$$

where $\xi(k) \in \mathbb{R}^n$ is the state vector of the dual system, $u(k) \in \mathbb{R}^p$ is the control input vector, and the initial state $\xi(0)$ is a independent Gaussian random vector $\xi(0) \sim \mathcal{N}(0,\, W_f)$ with $W_f \in \mathbb{S}_+^n$. Define the quadratic cost function

$$J_{\pi_N} := \mathrm{E}\left(\xi(N)^TQ_f\xi(N) + \sum_{k=0}^{N-1}\left(\xi(k)^TQ\xi(k) + u(k)^TRu(k)\right)\right)$$

over $\pi_N := (\mu_0,\, \mu_1, \ldots,\, \mu_{N-1})$ such that $u(k) = \mu_k(\mathbf{I}_k)$, where

$$\mathbf{I}_k := (\xi(0),\, \xi(1), \ldots,\, \xi(k),\, u(0),\, u(1), \ldots,\, u(k-1)).$$

Problem 1 (Stochastic LQR problem for the dual system)**.** *Solve*

$$\pi_N^* := \underset{\pi_N \in \Pi_N}{\operatorname{argmin}}\ J_{\pi_N} \quad \text{subjectto (6)}$$

where Π_N is the set of all admissible policies.

From the standard results of the stochastic LQR theory [1, page 150], the optimal solution is obtained as

$$\begin{aligned} &u(k) = F_k\xi(k), \quad F_k = L_{N-k-1}^T = -(R + CP_{N-k-1}C^T)^{-1}CP_{N-k-1}A^T \\ &k \in \{0,\, 1, \ldots,\, N-1\} \end{aligned} \tag{7}$$

and the optimal value of the cost function is $\mathrm{Tr}(W_fP_N)$.

III. Duality for analysis

In this section, we assume that the state-feedback gains F_k, $k \in \{0, 1, \dots, N-1\}$ are arbitrarily fixed, and then consider the following problem.

Problem 2. *Consider the dual system* (6) *and assume that* F_k, $k \in \{0, 1, \dots, N-1\}$ *are arbitrarily fixed. Compute the cost function value* J_{π_N} *with* $\pi_N = \{F_k \xi(k)\}_{k=0}^{N-1}$.

Proposition 1. *Let* $W_f \in \mathbb{S}_+^n$ *be given and consider the optimization problem*

$$\min_{S_1,\dots,S_N \in \mathbb{S}^n} \left(\mathrm{Tr}(S_N Q_f) + \sum_{k=0}^{N-1} \mathrm{Tr}(S_k Q + S_k L_{N-k-1} R L_{N-k-1}^T) \right) \tag{8}$$

subject to

$$S_{k+1} = (A^T + C^T F_k) S_k (A^T + C^T F_k)^T, \quad k \in \{0, 1, \dots, N-1\},$$

where $S_0 = W_f$*. The optimal objective function value of* (8) *is equal to the cost function value* J_{π_N} *in* Problem 2.

Proof. First of all, since

$$\begin{aligned}\mathrm{E}(\xi(k+1)\xi(k+1)^T) &= \mathrm{E}([A^T\xi(k) + C^T F_k \xi(k)][A^T\xi(k) + C^T F_k \xi(k)]^T) \\ &= (A^T + C^T F_k)\mathrm{E}(\xi(k)\xi(k)^T)(A^T + C^T F_k)^T,\end{aligned}$$

the covariance update equation of (6) is

$$S_{k+1} = (A^T + C^T F_k) S_k (A^T + C^T F_k)^T, \quad k \in \{0, 1, \dots, N-1\}, \quad S_0 = W_f,$$

where $S_k := \mathrm{E}(\xi(k)\xi(k)^T)$. Moreover, J_{π_N} can be written as

$$J_{\pi_N} = \mathrm{Tr}(S_N Q_f) + \sum_{k=0}^{N-1} \mathrm{Tr}(S_k Q + S_k L_{N-k-1} R L_{N-k-1}^T)$$

From the identities, Problem 2 is equivalent to the optimization 8. This completes the proof. $\square$

Remark 1. *The optimization* (8) *is a equality constrained optimization (linear programming problem) with a unique feasible point. Therefore, its optimal point is the unique feasible point.*

The dual problem of (8) is established in the following result.

Proposition 2. *Let $Q_f \in \mathbb{S}^n_+$ be given. The Lagrangian dual problem of (8) is given by*

$$\max_{P_1,\ldots,P_N \in \mathbb{S}^n} \operatorname{Tr}(S_0 P_N) \quad \text{subject to} \tag{9}$$
$$P_{k+1} = (A + L_k C)P_k(A + L_k C)^T + Q + L_k R L_k^T, \quad k \in \{0,\, 1, \ldots,\, N-1\},$$

where $P_0 = Q_f$ and $L_k = K_{N-k-1}^T$.

Proof. Introduce the Lagrangian for the optimization problem (8)

$$\mathcal{L}(\mathbf{S},\, \mathbf{P}) := \operatorname{Tr}(S_N Q_f) + \sum_{k=0}^{N-1} \operatorname{Tr}(S_k Q + S_k L_{N-k-1} R L_{N-k-1}^T)$$
$$+ \sum_{k=0}^{N-1} \operatorname{Tr}(P_{N-k-1}[(A + L_{N-k-1}C)^T \times S_k(A + L_{N-k-1}C) - S_{k+1}]),$$

where $\mathbf{S} := \{S_k\}_{k=1}^N$ and $\mathbf{P} := \{P_k\}_{k=1}^N$. The Lagrangian function $\mathcal{L}(\mathbf{S},\, \mathbf{P})$ can be written by

$$\mathcal{L}(\mathbf{S},\, \mathbf{P}) = \operatorname{Tr}((Q_f - P_0)S_N) + \operatorname{Tr}(P_N S_0)$$
$$+ \sum_{k=0}^{N-1} \operatorname{Tr}([(A + BF_k)^T P_{N-k-1}(A + BF_k) - P_{N-k} + Q + F_k^T R F_k] S_k).$$

The dual function is $\mathcal{D}(\mathbf{P}) = \inf_{\mathbf{S} \subset \mathbb{S}^n} \mathcal{L}(\mathbf{S},\, \mathbf{P})$, and the dual problem is $\sup_{\mathbf{P} \subset \mathbb{S}^n} \mathcal{D}(\mathbf{P})$. Since $\inf_{\mathbf{S}} \mathcal{L}(\mathbf{S},\, \mathbf{P})$ is finite only when the constraints in (9) are satisfied, the dual problem can be formulated as (9). For the unique dual feasible point $\mathbf{P} = \{P_k\}_{k=1}^N$ satisfying the constraints in (9), we have $\mathcal{L}(\mathbf{S},\, \mathbf{P}) = \operatorname{Tr}(P_N S_0)$. In addition, by a direct calculation, it can be proved that

$$\operatorname{Tr}(P_N S_0) = \sum_{k=0}^{N-1} \operatorname{Tr}(S_k Q + S_k L_{N-k-1} R L_{N-k-1}^T).$$

Since the objective function value of the dual feasible point and the objective function value of the primal feasible point are identical, both points are primal and dual optimal points, and there is no duality gap. This completes the proof. □

Remark 2. *Several remarks are in order.*

1) *The result of Proposition 2 can be also obtained using algebraic manipulations (without using the Lagrangian duality).*
2) *The constraints in (9) are equivalent to the Riccati equation (4). Therefore, the matrices $P_1, \ldots,\, P_N$ can be interpreted as the Lagrange multipliers for the equality constraints in (8).*

3) The constraints in (9) can be viewed as a covariance update of the estimation error system (5).

4) The cost function value of the terminal error of (5) has the same value as the quadratic cost function value (5) of the dual system. Roughly speaking, the existence of the quadratic cost function corresponds to the existence of the noises in its dual system.

Conversely, consider the following problem.

Problem 3. *Assume that the estimator gains L_k, $k \in \{0, 1, \ldots, N-1\}$ are arbitrarily fixed. For the estimation error system (5), compute the cost function value*

$$J_{\text{ob}} := \mathrm{E}\left(e(N)^T W_f e(N) + \sum_{k=0}^{N-1} e(k)^T W e(k)\right) \tag{10}$$

Problem 3 can be converted into the covariance optimization problem.

Proposition 3. *Consider the optimization problem*

$$\min_{P_1,\ldots,P_N \in \mathbb{S}^n} \mathrm{Tr}(P_N W_f) + \sum_{k=0}^{N-1} \mathrm{Tr}(P_k W) \tag{11}$$

$$\text{subject to}$$

$$P_{k+1} = (A + L_k C) P_k (A + L_k C)^T + Q + L_k R L_k^T, k \in \{0, 1, \ldots, N-1\},$$

where $P_0 = Q_f$. The optimal objective function value of (11) is equal to the cost function value J_{ob} in Problem 3.

Proof. Straightforward from the previous results. □

This problem has a unique feasible point, and the matrix equality constraints are the covariance updates of the estimation error system. Following similar lines to the proof of Proposition 2, its Lagrangian dual problem can be obtained.

Proposition 4. *Assume that the estimator gains L_k, $k \in \{0, 1, \ldots, N-1\}$ are arbitrarily fixed. The Lagrangian dual problem of (11) is given by*

$$\max_{S_1,\ldots,S_N \in \mathbb{S}^n} \mathrm{Tr}(Q_f S_N) + \sum_{k=0}^{N-1} \mathrm{Tr}([Q + L_k R L_k^T] S_{N-k1}) \tag{12}$$

$$\text{subject to}$$

$$S_{k+1} = (A + L_{N-k-1} C)^T S_k (A + L_{N-k-1} C) + W, \quad k \in \{0, 1, \ldots, N-1\},$$

where $S_0 = W_f$.

Proof. It can be readily proved following similar lines to the proof of Proposition 2. $\square$

Remark 3. *The matrix equality constraints of* (12) *can be interpreted as the covariance update of the dual system*

$$\xi(k+1) = (A^T + C^T L_{N-k-1}^T)\xi(k) + \sigma(k) \tag{13}$$

where $\xi(0) \sim \mathcal{N}(0,\, W_f)$ *and* $\sigma(k) \sim \mathcal{N}(0,\, W)$ *are independent Gaussian random vectors. The objective function of* (12) *can be also written by* J_{π_N} *with* $\pi_N = \{L_{N-k-1}\xi(0)\}_{k=0}^{N-1}$. *Therefore, we have* $J_{\text{ob}} = J_{\pi_N}$.

IV. Kalman filtering problem in the covariance optimization form

In this section, we will study the Kalman filtering problem in the covariance optimization form, and discuss about its solution. Consider the estimation error system (5) and the corresponding quadratic cost function (10) again.

Problem 4. *Solve* $\min\limits_{L_0,\ldots,\, L_{N-1} \in \mathbb{R}^{n\times m}} J_{\text{ob}}$.

From the results of the previous section, it can be proved that Problem 4 is equivalent to the following covariance optimization problem.

Problem 5. *Solve*

$$\min_{\substack{P_1,\ldots,\, P_N \in \mathbb{S}^n \\ L_0,\ldots,\, L_{N-1} \in \mathbb{R}^{n\times m}}} \mathrm{Tr}(P_N W_f) + \sum_{k=0}^{N-1} \mathrm{Tr}(P_k W)$$

$$\text{subject to}$$

$$P_{k+1} = (A + L_k C) P_k (A + L_k C)^T + Q + L_k R L_k^T, \quad k \in \{0,\, 1,\ldots,\, N-1\},$$

with $P_0 = Q_f$.

Regarding this problem, we can make the following conclusions.

Proposition 5. *Let* $\{L_k^*,\, P_{k+1}^*\}_{k=0}^{N-1}$ *be an optimal solution to* Problem 5. *Then, it is equivalent to the pairs of the Kalman gain matrices* (3) *and the corresponding covariance matrices* (4), *respectively.*

Proof. By plugging each P_k in the equality constraints of Problem 5 into its objective function, Problem 5 can be written by the unconstrained optimization problem

$$\min_{L_0,\dots,\,L_{N-1}\in\mathbb{R}^{n\times m}} \quad \Gamma(\{L_i\}_{i=0}^{N-1})$$

with some function Γ. By algebraic manipulations, it can be proved that, for each L_k, $\Gamma(\{L_i\}_{i=0}^{N-1})$ can be written as

$$\Gamma(\{L_i\}_{i=0}^{N-1}) = \mathrm{Tr}([(A+L_kC)P_k(A+L_kC)^T + Q + L_kRL_k^T]M_k) + \gamma_k,$$

for some $M_k \in \mathbb{S}_+^n$ and $\gamma_k > 0$. Rearranging terms, it can be rewritten as

$$\begin{aligned}
&\Gamma(\{L_i\}_{i=0}^{N-1}) = \mathrm{Tr}([L_k[CP_kC^T + R]L_k^T]M_k)\\
&+ 2\mathrm{Tr}(L_kCP_kA^TM_k) + \mathrm{Tr}(AP_kA^TM_k) + \mathrm{Tr}(QM_k) + \gamma_k\\
&= \mathrm{Tr}(\mathrm{vec}(L_k^T)^T(M_k \otimes [CP_kC^T + R])\mathrm{vec}(L_k^T))\\
&+ 2\mathrm{Tr}(L_kCP_kA^TM_k) + \mathrm{Tr}(AP_kA^TM_k) + \mathrm{Tr}(QM_k) + \gamma_k.
\end{aligned}$$

Since $M_k \otimes [CP_kC^T + R] \in \mathbb{S}_+^n$, $\Gamma(\{L_i\}_{i=0}^{N-1})$ is a convex quadratic function with respect to each L_k. Therefore, setting $\partial f(L_k)/\partial L_k = 0$, $\bar{L}_k \in \mathrm{argmin}_{L_k\in\mathbb{R}^{n\times m}}\Gamma(\{L_i\}_{i=0}^{N-1})$ is obtained as $\bar{L}_k = -AP_kC^T(R + CP_kC^T)^{-1}$. Since the optimal solution $\bar{L}_k$ does not depend on $\{L_{N-1},\dots,\,L_k\}$, it can be proved that the principle of optimality holds. In other words, if we define the function

$$V_t(\{L_{N-1},\dots,\,L_t\}) := \min_{L_0,\dots,\,L_{t-1}\in\mathbb{R}^{n\times m}} \Gamma(\{L_{N-1},\dots,\,L_t,\,L_{t-1},\dots,\,L_0\})$$

for all $t \in \{1,\,2,\dots,\,N-1\}$, then it obeys

$$V_{t+1}(L_{N-1},\dots,L_{t+1}) = \min_{L_t\in\mathbb{R}^{n\times m}} V_t(L_{N-1},\dots,L_t)$$

for all $t \in \{1,\,2,\dots,\,N-1\}$. Therefore, a global minimizer of Problem 5 can be obtained sequentially by solving

$$\begin{aligned}
&\bar{L}_k = \underset{L_k\in\mathbb{R}^{n\times m}}{\mathrm{argmin}}\ \mathrm{Tr}(P_{k+1})\\
&\text{subject to}\\
&P_{k+1} = (A+L_kC)\bar{P}_k(A+L_kC)^T + Q + L_kRL_k^T
\end{aligned}$$

for all $k \in \{0,\,1,\dots,\,N-1\}$. Since $\bar{P}_k \in \mathbb{S}_{++}^n$, each subproblem is a convex quadratic programming with a unique solution $\bar{L}_k = -AP_kC^T(R+CP_kC^T)^{-1}$. Therefore, a global solution to Problem 5 is identical to the solution to the Kalman filtering problem (3). This completes the proof. □

V. DUALITY OF THE LQR PROBLEM

Consider the stochastic LTI system

$$x(k+1) = Ax(k) + Bu(k) + w(k) \tag{14}$$

where $k \in \mathbb{N}$, $x(k) \in \mathbb{R}^n$ is the state vector, $u(k) \in \mathbb{R}^m$ is the control input vector, $x(0) \sim \mathcal{N}(0, W_f)$ and $w(k) \sim \mathcal{N}(0, W)$ are independent Gaussian random vectors. Define the quadratic cost function of (14)

$$J_{\pi_N} := \mathrm{E}\left(x(N)^T Q_f x(N) + \sum_{k=0}^{N-1} (x(k)^T Q x(k) + u(k)^T R u(k)) \right)$$

over $\pi_N := (\mu_0, \mu_1, \ldots, \mu_{N-1})$ such that $u(k) = \mu_k(\mathbf{I}_k)$, where

$$\mathbf{I}_k := (x(0), x(1), \ldots, x(k), u(0), u(1), \ldots, u(k-1))$$

Then, the stochastic LQR problem can be stated as follows.

Problem 6 (Stochastic LQR problem)**.** *Solve*

$$\pi_N^* := \operatorname*{argmin}_{\pi_N \in \Pi_N} J_{\pi_N} \quad \text{subject to (14)}$$

where Π_N is the set of all admissible policies.

The goal of this section is to derive a dual form of the stochastic LQR problem. First of all, the stochastic LQR problem can be formulated as the following covariance optimization problem.

Problem 7. *Solve*

$$\min_{\substack{F_0,\ldots,F_{N-1} \in \mathbb{R}^{m\times n} \\ S_1,\ldots,S_N \in \mathbb{S}^n}} \mathrm{Tr}(Q_f S_N) + \sum_{k=0}^{N-1} \mathrm{Tr}([Q + F_k^T R F_k] S_k)$$

$$\text{subject to}$$

$$S_{k+1} = (A + BF_k) S_k (A + BF_k)^T + W, \quad k \in \{0, 1, \ldots, N-1\},$$

where $S_0 = W_f$.

The Lagrange dual problem of Problem 7 is established below.

Proposition 6. *The Lagrange dual problem of Problem 7 is given by*

$$\max_{P_1,\ldots,P_N \in \mathbb{S}^n} \mathrm{Tr}(P_N W_f) + \sum_{k=0}^{N-1} \mathrm{Tr}(P_k W) \tag{15}$$

subject to

$$P_{k+1} = A^T P_k A - A^T P_k B(R + B^T P_k B)^{-1} B^T P_k A + Q, \quad k \in \{0, 1, \dots, N-1\}.$$

with $P_0 = Q_f$*. In addition, the strong duality holds, and the primal optimal point* $\{S_k, F_{k-1}\}_{k=1}^{N}$ *and dual optimal point* $\{P_k\}_{k=0}^{N-1}$ *are given by*

$$\begin{aligned}
&F_k = -(R + B^T P_{N-k-1} B)^{-1} B^T P_{N-k-1} A \\
&S_{k+1} = (A + F_k B) S_k (A + F_k B)^T + W \\
&P_{k+1} = A^T P_k A - A^T P_k B(R + B^T P_k B)^{-1} B^T P_k A + Q \\
&k \in \{0, 1, \dots, N-1\}
\end{aligned}$$

Proof. Define the Lagrangian function of Problem 7

$$\begin{aligned}
&\mathcal{L}(\mathbf{S}, \mathbf{F}, \mathbf{P}) := \mathrm{Tr}(S_N Q_f) + \sum_{k=0}^{N-1} \mathrm{Tr}(S_k Q + S_k F_k^T R F_k) \\
&+ \sum_{k=0}^{N-1} \mathrm{Tr}([(A + BF_k) S_k (A + BF_k)^T + W - S_{k+1}] P_{N-k-1}),
\end{aligned}$$

where $\mathbf{S} := \{S_k\}_{k=1}^{N}, \mathbf{F} := \{F_k\}_{k=0}^{N-1}, \mathbf{P} := \{P_k\}_{k=0}^{N-1}$. Rearranging some terms, it can be represented by

$$\begin{aligned}
&\mathcal{L}(\mathbf{S}, \mathbf{F}, \mathbf{P}) = \mathrm{Tr}((Q_f - P_0) S_N) + \mathrm{Tr}(P_N S_0) \\
&+ \sum_{k=0}^{N-1} \mathrm{Tr}([(A + BF_k)^T P_{N-k-1} (A + BF_k) - P_{N-k} + Q + F_k^T R F_k] S_k)
\end{aligned}$$

The dual function is $\mathcal{D}(\mathbf{P}) := \inf_{\mathbf{F}, \mathbf{S}} \mathcal{L}(\mathbf{S}, \mathbf{F}, \mathbf{P})$, and the Lagrangian dual problem is $\sup_{\mathbf{P}} \mathcal{D}(\mathbf{P})$. We first prove that the Lagrangian function $\mathcal{L}(\mathbf{S}, \mathbf{F}, \mathbf{P})$ is convex in $\mathbf{S}, \mathbf{F}$ under a certain condition on $\mathbf{P}$. Since $S_k \succ 0$, $k \in \{1, 2, \dots, N\}$, there exist nonsingular matrices Z_k, $k \in \{1, 2, \dots, N\}$ such that $S_k = Z_k Z_k^T$, $k \in \{1, 2, \dots, N\}$. Letting $F_k Z_k = G_k$, it can be proved that $\inf_{\mathbf{F}, \mathbf{S}} \mathcal{L}(\mathbf{S}, \mathbf{F}, \mathbf{P})$ is equivalent to minimizing the following function with respect to Z_k and G_k:

$$\begin{aligned}
&\mathrm{Tr}(Z_N^T Q_f Z_N) \\
&+ \sum_{k=0}^{N-1} \mathrm{Tr}(Z_k^T Q Z_k + G_k^T R G_k) \\
&+ \sum_{k=0}^{N-1} \mathrm{Tr}((AZ_k + BG_k)^T P_{N-k-1} (AZ_k + BG_k) + W P_{N-k-1} - Z_{k+1}^T P_{N-k-1} Z_{k+1})
\end{aligned}$$

$$
\begin{aligned}
&= \mathrm{vec}(Z_N)^T (I \otimes (Q_f - P_0)) \mathrm{vec}(Z_N) \\
&+ \begin{bmatrix} \mathrm{vec}(Z_0) \\ \mathrm{vec}(G_0) \end{bmatrix}^T \Phi \begin{bmatrix} \mathrm{vec}(Z_0) \\ \mathrm{vec}(G_0) \end{bmatrix} + \sum_{k=1}^{N-1} \begin{bmatrix} \mathrm{vec}(Z_k) \\ \mathrm{vec}(G_k) \end{bmatrix}^T \Omega_k \begin{bmatrix} \mathrm{vec}(Z_k) \\ \mathrm{vec}(G_k) \end{bmatrix} + \sum_{k=1}^{N-1} W P_{N-k-1}
\end{aligned}
$$

where

$$
\Phi := \begin{bmatrix} I \otimes (Q + A^T P_{N-1} A) & I \otimes A^T P_{N-1} B \\ I \otimes B^T P_{N-1} A & I \otimes (R + B^T P_{N-1} B) \end{bmatrix},
$$

$$
\Omega_k := \begin{bmatrix} I \otimes (Q + A^T P_{N-k-1} A - P_{N-k}) & I \otimes A^T P_{N-k-1} B \\ I \otimes B^T P_{N-k-1} A & I \otimes (R + B^T P_{N-k-1} B) \end{bmatrix}
$$

The above function is quadratic. If $Q_f - P_0 \succeq 0$ and $Q + A^T P_{N-k-1} A - P_{N-k} \succeq 0$ for each k, then it is convex. Otherwise, we have $\inf_{\mathbf{F},\mathbf{S}} \mathcal{L}(\mathbf{S}, \mathbf{F}, \mathbf{P}) = -\infty$, which implies that the given $\mathbf{P}$ should not be dual feasible. With $\mathbf{P}$ satisfying the two conditions, solving $\inf_{\mathbf{F},\mathbf{S}} \mathcal{L}(\mathbf{S}, \mathbf{F}, \mathbf{P})$ is a convex optimization problem. Letting the derivatives of $\mathcal{L}(\mathbf{S}, \mathbf{F}, \mathbf{P})$ with respect to $\{\mathbf{S}, \mathbf{F}\}$ be zero, a primal feasible point $(\mathbf{S}^*, \mathbf{F}^*) = \inf_{\mathbf{F},\mathbf{S}} \mathcal{L}(\mathbf{S}, \mathbf{F}, \mathbf{P})$ can be obtained as

$$
\begin{aligned}
&F_k = -R + B^T P_{N-k-1} B)^{-1} B^T P_{N-k-1} A \\
&S_{k+1} = (A + F_k B) S_k (A + F_k B)^T + W \\
&S_0 = W_f, \quad k \in \{0, 1, \ldots, N-1\}
\end{aligned}
$$

Let $\mathbf{F}^* := \{-(R + B^T P_{N-k-1} B)^{-1} B^T P_{N-k-1} A\}_{k=0}^{N-1}$. Since $\mathbf{F}^*$ is not dependent on S_k, $k \in \{1, 2, \ldots, N\}$, we first plug $\mathbf{F}^*$ into $\mathcal{L}(\mathbf{S}, \mathbf{F}, \mathbf{P})$ and obtain

$$
\begin{aligned}
&\mathcal{L}(\mathbf{S}, \mathbf{F}^*, \mathbf{P}) = \mathrm{Tr}((Q_f - P_0) S_N) + \mathrm{Tr}(P_N S_0) + \sum_{k=0}^{N-1} \mathrm{Tr}(P_{N-k-1} W) \\
&+ \sum_{k=0}^{N-1} \mathrm{Tr}([A^T P_{N-k-1} A - A^T P_{N-k-1} B (R + B^T P_{N-k-1} B)^{-1} B^T P_{N-k-1} A + Q - P_{N-k}] S_k)
\end{aligned}
$$

Then, the dual function $\mathcal{D}(\mathbf{P}) := \inf_{\mathbf{F},\mathbf{S}} \mathcal{L}(\mathbf{S}, \mathbf{F}, \mathbf{P}) = \inf_{\mathbf{S}} \mathcal{L}(\mathbf{S}, \mathbf{F}^*, \mathbf{P})$ has a finite value only when the constraints in (15) hold. Therefore, the dual problem $\sup_{\mathbf{P}} \mathcal{D}(\mathbf{P})$ is given by (15). Finally, note that the dual problem has a unique feasible point, and this implies that the dual feasible point is also the dual optimal point. By plugging it into its objective and rearranging terms, we can prove that the dual objective function value is the same as the primal objective function value. By the weak duality, the objective value of the primal optimal point should be larger than or equal to the objective value of the dual optimal point. Since both objective values are identical, there is no duality gap, and $\mathbf{F}^*$ is the primal optimal point. This completes the proof. □

Remark 4. *The result of Proposition 6 proves that the Riccati equation and its solution corresponds to the equality constraints of the Lagrange dual problem and the Lagrange multipliers, respectively, of Problem 7.*

VI. Conclusion

In this paper, the relation between the Kalman filtering and the LQR problems was studied by using the Lagrangian duality theory. We first arbitrary fixed the gain matrices of the Kalman filtering and proved that the Kalman filtering problem is a Lagrangian dual problem of the LQR problem. Next, we considered the case that the Kalman gain matrices are also the optimization parameters. In this case, the problem becomes harder because it is not clear whether or not the optimization formulation of the Kalman filtering problem is convex. It is proved that the optimal solution to the optimization formulation of the Kalman filtering problem can be derived as the standard Ricatti equations.

On the other hand, the Lagrangian dual problem of the LQR problem was derived as well. It was proved that the solution of the Ricatti equation is the optimal solution to the Lagrangian dual problem of the LQR problem.

In the future work, a clearer connection between the Kalman filtering and the LQR problem will be explored.

References

[1] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, 3rd ed. Nashua, MA: Athena Scientific, 2005, vol. 1.

[2] M. Green and D. J. Limebeer, *Linear robust control*. Courier Corporation, 2012.

[3] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[4] D. Henrion, G. Meinsma *et al.*, "Rank-one LMIs and Lyapunov's inequality," *IEEE Transactions on Automatic Control*, vol. 46, no. 8, pp. 1285–1288, 2001.

[5] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM review*, vol. 38, no. 1, pp. 49–95, 1996.

[6] D. D. Yao, S. Zhang, and X. Y. Zhou, "Stochastic linear-quadratic control via semidefinite programming," *SIAM Journal on Control and Optimization*, vol. 40, no. 3, pp. 801–823, 2001.

[7] M. A. Rami and X. Y. Zhou, "Linear matrix inequalities, Riccati equations, and indefinite stochastic linear quadratic controls," *Automatic Control, IEEE Transactions on*, vol. 45, no. 6, pp. 1131–1143, 2000.

[8] V. Balakrishnan and L. Vandenberghe, "Semidefinite programming duality and linear time-invariant systems," *Automatic Control, IEEE Transactions on*, vol. 48, no. 1, pp. 30–41, 2003.

[9] A. Gattami, “Generalized linear quadratic control,” *IEEE Transactions on Automatic Control*, vol. 55, no. 1, pp. 131–136, 2010.

[10] S. You and J. C. Doyle, “A Lagrangian dual approach to the Generalized KYP lemma,” in *CDC*, 2013, pp. 2447–2452.