# 10

# Duality Between Constrained Estimation and Control

## 10.1 Overview

The previous chapter showed that the problem of constrained estimation can be formulated as a constrained optimisation problem. Indeed, this problem is remarkably similar to the constrained control problem—differing only with respect to the boundary conditions. (In control, the initial condition is fixed, whereas in estimation, the initial condition can also be adjusted.) In the current chapter we show that the similarity between the two problems of constrained estimation and constrained control has deeper implications.

In particular, we derive the Lagrangian dual (see Section 2.6 of Chapter 2) of a constrained estimation problem and show that it leads to a particular unconstrained nonlinear optimal control problem. We then show that the original (primal) constrained estimation problem has an equivalent formulation as an unconstrained nonlinear optimisation problem, exposing a clear symmetry with its dual.

## 10.2 Lagrangian Duality of Constrained Estimation and Control

Consider the following system

$$\begin{aligned} x_{k+1} &= Ax_k + Bw_k \quad \text{for } k = 0, \cdots, N-1, \\ y_k &= Cx_k + v_k \quad \text{for } k = 1, \cdots, N, \end{aligned} \tag{10.1}$$

where $x_k \in \mathbb{R}^n$, $w_k \in \mathbb{R}^m$, $y_k \in \mathbb{R}^p$. For clarity of exposition, we begin with the case where only the process noise sequence $\{w_k\}$ is constrained.[1] We thus assume that $\{w_k\}$ is an i.i.d. sequence having truncated Gaussian distribution

[1] The case of general constraints on $w_k$, $v_k$ and $x_0$ will be treated in Sections 10.6 and 10.7.

of the form given in (9.11) of Chapter 9, with $\Omega_1 = \Omega$. We further assume that $\{v_k\}$ is an i.i.d. sequence having a Gaussian distribution $N(0, R)$, and $x_0$ has a Gaussian distribution $N(\mu_0, P_0)$.

For (10.1) we consider the optimisation problem defined in (9.29)–(9.35) of Chapter 9, which yields the joint a posteriori most probable state estimates. According to the assumptions, we set $\Omega_1 = \Omega$, $\Omega_2 = \mathbb{R}^p$, and $\Omega_3 = \mathbb{R}^n$ in (9.29)–(9.35). Thus, we consider:

$$\mathcal{P}_e: \quad V_N^{\text{OPT}}(\mu_0, \{y_k^d\}) \triangleq \min_{\hat{x}_k, \hat{v}_k, \hat{w}_k} V_N(\{\hat{x}_k\}, \{\hat{v}_k\}, \{\hat{w}_k\}), \tag{10.2}$$

subject to:

$$\hat{x}_{k+1} = A\hat{x}_k + B\hat{w}_k \quad \text{for } k = 0, \ldots, N-1, \tag{10.3}$$

$$\hat{v}_k = y_k^d - C\hat{x}_k \quad \text{for } k = 1, \ldots, N, \tag{10.4}$$

$$\{\hat{x}_0, \ldots, \hat{x}_N, \hat{v}_1, \ldots, \hat{v}_N, \hat{w}_0, \ldots, \hat{w}_{N-1}\} \in X, \tag{10.5}$$

where, in (10.5),

$$X = \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{N+1} \times \underbrace{\mathbb{R}^p \times \cdots \times \mathbb{R}^p}_{N} \times \underbrace{\Omega \times \cdots \times \Omega}_{N}. \tag{10.6}$$

In (10.2), the objective function is

$$V_N(\{\hat{x}_k\}, \{\hat{v}_k\}, \{\hat{w}_k\}) \triangleq \frac{1}{2}(\hat{x}_0 - \mu_0)^{\mathrm{T}} P_0^{-1} (\hat{x}_0 - \mu_0) + \frac{1}{2}\sum_{k=0}^{N-1} \hat{w}_k^{\mathrm{T}} Q^{-1} \hat{w}_k + \frac{1}{2}\sum_{k=1}^{N} \hat{v}_k^{\mathrm{T}} R^{-1} \hat{v}_k, \tag{10.7}$$

where $P_0 > 0$, $Q > 0$, $R > 0$ are the covariance matrices in (9.11)–(9.13).

The following result establishes duality between the constrained *estimation* problem $\mathcal{P}_e$ and a particular unconstrained nonlinear optimal *control* problem.

**Theorem 10.2.1 (Dual Problem)** *Assume $\Omega$ in* (10.6) *is a nonempty closed convex set. Given the primal constrained fixed horizon estimation problem $\mathcal{P}_e$ defined by equations* (10.2)–(10.7)*, the Lagrangian dual problem is*

$$\mathcal{D}_e: \quad \phi^{\text{OPT}}(\mu_0, \{y_k^d\}) \triangleq \min_{\lambda_k, u_k} \phi(\{\lambda_k\}, \{u_k\}), \tag{10.8}$$

*subject to:*

$$\lambda_{k-1} = A^{\mathrm{T}}\lambda_k + C^{\mathrm{T}} u_k \quad \textit{for } k = 1, \cdots, N, \tag{10.9}$$

$$\lambda_N = 0, \tag{10.10}$$

$$\zeta_k = B^{\mathrm{T}}\lambda_k \quad \textit{for } k = 0, \cdots, N-1, \tag{10.11}$$

$$\bar{\zeta}_k = Q^{-1/2}\Pi_{\tilde{\Omega}} Q^{1/2}\zeta_k \quad \textit{for } k = 0, \cdots, N-1. \tag{10.12}$$

*In* (10.8)*, the objective function is*

$$\begin{aligned}\phi(\{\lambda_k\},\{u_k\}) \triangleq \frac{1}{2}(A^{\mathrm{T}}\lambda_0 + P_0^{-1}\mu_0)^{\mathrm{T}} P_0 (A^{\mathrm{T}}\lambda_0 + P_0^{-1}\mu_0) \\ + \frac{1}{2}\sum_{k=1}^{N}(u_k - R^{-1}y_k^d)^{\mathrm{T}} R (u_k - R^{-1}y_k^d) \\ + \sum_{k=0}^{N-1}\left[\frac{1}{2}\bar{\zeta}_k^{\mathrm{T}} Q\bar{\zeta}_k + (\zeta_k - \bar{\zeta}_k)^{\mathrm{T}} Q\bar{\zeta}_k\right] + \gamma \end{aligned} \tag{10.13}$$

*where $\gamma$ is the constant term given by*

$$\gamma \triangleq -\frac{1}{2}\mu_0^{\mathrm{T}} P_0^{-1}\mu_0 - \frac{1}{2}\sum_{k=1}^{N}(y_k^d)^{\mathrm{T}} R^{-1} y_k^d. \tag{10.14}$$

*In* (10.12), $\Pi_{\tilde{\Omega}}$ *denotes the minimum Euclidean distance projection onto* $\tilde{\Omega} \triangleq \{z : Q^{1/2}z \in \Omega\}$, *that is,*

$$\begin{aligned}\Pi_{\tilde{\Omega}} : \mathbb{R}^m &\longrightarrow \tilde{\Omega} \\ s &\longmapsto \bar{s} = \Pi_{\tilde{\Omega}} s \triangleq \arg\min_{z\in\tilde{\Omega}} \|z - s\|.\end{aligned} \tag{10.15}$$

*Moreover, there is no duality gap, that is, the minimum achieved in* (10.2) *is equal to minus the minimum achieved in* (10.8).

*Proof.* Consider the primal constrained fixed horizon estimation problem $\mathcal{P}_e$, defined by equations (10.2)–(10.7). From (2.44) in Chapter 2, the Lagrangian dual function $\theta$ is given by:

$$\theta\big(\{\lambda_k\},\{u_k\}\big) = \inf_{\hat{w}_k\in\Omega,\hat{x}_k,\hat{v}_k} L\big(\{\hat{x}_k\},\{\hat{v}_k\},\{\hat{w}_k\},\{\lambda_k\},\{u_k\}\big), \tag{10.16}$$

where the function $L$ is defined as,

$$\begin{aligned} L\big(\{\hat{x}_k\},\{\hat{v}_k\},\{\hat{w}_k\},\{\lambda_k\},\{u_k\}\big) = V_N(\{\hat{x}_k\},\{\hat{v}_k\},\{\hat{w}_k\}) \\ + \sum_{k=0}^{N-1}\lambda_k^{\mathrm{T}}\Big[\hat{x}_{k+1} - A\hat{x}_k - B\hat{w}_k\Big] \\ + \sum_{k=1}^{N} u_k^{\mathrm{T}}\Big[y_k^d - C\hat{x}_k - \hat{v}_k\Big]. \end{aligned} \tag{10.17}$$

In (10.17), $V_N$ is the primal objective function defined in (10.7), and $\{\lambda_k\}$ and $\{u_k\}$ are the Lagrange multipliers corresponding, respectively, to the linear equalities (10.3) and (10.4). Using (10.7) in (10.17), and combining terms, the function $L$ can be rewritten as

$$
\begin{aligned}
L\big(\{\hat{x}_k\},\{\hat{v}_k\},\{\hat{w}_k\},\{\lambda_k\},\{u_k\}\big) &= \frac{1}{2}(\hat{x}_0-\mu_0)^{\mathrm{T}}P_0^{-1}(\hat{x}_0-\mu_0)-\lambda_0^{\mathrm{T}}A\hat{x}_0\\
&\quad+\sum_{k=1}^{N}\Big\{\frac{1}{2}\hat{v}_k^{\mathrm{T}}R^{-1}\hat{v}_k-u_k^{\mathrm{T}}\hat{v}_k+u_k^{\mathrm{T}}y_k^d\Big\}\\
&\quad+\sum_{k=0}^{N-1}\Big\{\frac{1}{2}\hat{w}_k^{\mathrm{T}}Q^{-1}\hat{w}_k-\lambda_k^{\mathrm{T}}B\hat{w}_k\Big\}\\
&\quad+\sum_{k=1}^{N-1}\big\{(\lambda_{k-1}-A^{\mathrm{T}}\lambda_k-C^{\mathrm{T}}u_k)^{\mathrm{T}}\hat{x}_k\big\}\\
&\quad+(\lambda_{N-1}-C^{\mathrm{T}}u_N)^{\mathrm{T}}\hat{x}_N. \qquad (10.18)
\end{aligned}
$$

Notice that the terms that depend on the constrained variables $\hat{w}_k$ are independent of the other variables, $\hat{x}_k$ and $\hat{v}_k$, with respect to which the minimisation (10.16) is carried out. The values that achieve the infimum in (10.16), denoted $\hat{w}_k^*$, $\hat{x}_k^*$ and $\hat{v}_k^*$, can be computed from

$$\hat{w}_k^* = \arg\min_{\hat{w}_k\in\Omega}\Big\{\frac{1}{2}\hat{w}_k^{\mathrm{T}}Q^{-1}\hat{w}_k-\lambda_k^{\mathrm{T}}B\hat{w}_k\Big\} \quad \text{for } k=0,\cdots,N-1, \qquad (10.19)$$

$$\frac{\partial L(\cdot)}{\partial \hat{x}_0} = P_0^{-1}\,(\hat{x}_0^*-\mu_0)-A^{\mathrm{T}}\lambda_0=0, \qquad (10.20)$$

$$\frac{\partial L(\cdot)}{\partial \hat{v}_k} = R^{-1}\hat{v}_k^*-u_k=0 \quad \text{for } k=1,\cdots,N, \qquad (10.21)$$

provided that the following two conditions are satisfied

$$\lambda_{k-1}-A^{\mathrm{T}}\lambda_k-C^{\mathrm{T}}u_k=0 \quad \text{for } k=1,\cdots,N-1, \qquad (10.22)$$

$$\lambda_{N-1}-C^{\mathrm{T}}u_N=0. \qquad (10.23)$$

Notice from (10.18) that the infimum in (10.16) is $-\infty$ whenever $\{\lambda_k\}$ and $\{u_k\}$ are such that (10.22) and (10.23) are not satisfied. However, since we will subsequently choose $\{\lambda_k\}$ and $\{u_k\}$ so as to maximise $\theta(\{\lambda_k\},\{u_k\})$ in (10.16) (see (2.43) and (2.44) in Chapter 2), we are here interested only in those values of $\{\lambda_k\}$ and $\{u_k\}$ satisfying (10.22) and (10.23).

We next define the variables

$$\zeta_k \triangleq B^{\mathrm{T}}\lambda_k, \qquad (10.24)$$

$$s \triangleq Q^{-1/2}\hat{w}_k, \qquad (10.25)$$

$$s^* \triangleq Q^{-1/2}\hat{w}_k^*, \qquad (10.26)$$

which transform the minimisation problem (10.19) into the minimum Euclidean distance problem

$$s^* = \arg\min_{s\in\tilde{\Omega},}\Big\{\frac{1}{2}s^{\mathrm{T}}s-(\zeta_k^{\mathrm{T}}Q^{1/2})s\Big\}, \qquad (10.27)$$

where $\tilde{\Omega} \triangleq \{z : Q^{1/2}z \in \Omega\}$. The solution to (10.27) can be expressed as

$$s^* = \bar{s} \triangleq \Pi_{\tilde{\Omega}} Q^{1/2}\zeta_k, \tag{10.28}$$

where $\Pi_{\tilde{\Omega}}$ is the Euclidean projection (10.15). Using (10.26) and (10.28), the solution to (10.19) is then

$$\hat{w}_k^* = Q^{1/2}\Pi_{\tilde{\Omega}} Q^{1/2}\zeta_k. \tag{10.29}$$

Finally, we define

$$\bar{\zeta}_k \triangleq Q^{-1}\hat{w}_k^* = Q^{-1/2}\Pi_{\tilde{\Omega}} Q^{1/2}\zeta_k, \tag{10.30}$$

and introduce an extra variable, $\lambda_N \triangleq 0$, for ease of notation. Thus, from (10.19)–(10.24) and (10.30), we obtain:

$$\begin{aligned}
\hat{w}_k^* &= Q\bar{\zeta}_k \quad \text{for } k = 0, \cdots, N-1, &(10.31)\\
\bar{\zeta}_k &\triangleq Q^{-1/2}\Pi_{\tilde{\Omega}} Q^{1/2}\zeta_k \quad \text{for } k = 0, \cdots, N-1, &(10.32)\\
\zeta_k &\triangleq B^{\mathrm{T}}\lambda_k \quad \text{for } k = 0, \cdots, N-1, &(10.33)\\
\lambda_N &\triangleq 0, &(10.34)\\
\lambda_{k-1} &= A^{\mathrm{T}}\lambda_k + C^{\mathrm{T}}u_k \quad \text{for } k = 1, \cdots, N, &(10.35)\\
\hat{x}_0^* &= P_0 A^{\mathrm{T}}\lambda_0 + \mu_0, &(10.36)\\
\hat{v}_k^* &= Ru_k \quad \text{for } k = 1, \cdots, N. &(10.37)
\end{aligned}$$

Substituting (10.31)–(10.37) into (10.18) we obtain, after some algebraic manipulations, the Lagrangian dual function:

$$\begin{aligned}
\theta\big(\{\lambda_k\}, \{u_k\}\big) &= L\big(\{\hat{x}_k^*\}, \{\hat{v}_k^*\}, \{\hat{w}_k^*\}, \{\lambda_k\}, \{u_k\}\big)\\
&= -\frac{1}{2}\big\{\lambda_0^{\mathrm{T}} A P_0 A^{\mathrm{T}}\lambda_0 + 2\lambda_0^{\mathrm{T}} A\mu_0\big\}\\
&\quad - \frac{1}{2}\sum_{k=1}^{N}\big\{u_k^{\mathrm{T}} R u_k - 2u_k^{\mathrm{T}} y_k^d\big\}\\
&\quad + \sum_{k=0}^{N-1}\Big\{\frac{1}{2}\bar{\zeta}_k^{\mathrm{T}} Q\bar{\zeta}_k - \zeta_k^{\mathrm{T}} Q\bar{\zeta}_k\Big\}.
\end{aligned} \tag{10.38}$$

Finally, completing the squares in (10.38), and after further algebraic manipulations, we obtain:

$$\begin{aligned}
\theta\big(\{\lambda_k\}, \{u_k\}\big) &= -\frac{1}{2}(A^{\mathrm{T}}\lambda_0 + P_0^{-1}\mu_0)^{\mathrm{T}} P_0 (A^{\mathrm{T}}\lambda_0 + P_0^{-1}\mu_0)\\
&\quad - \frac{1}{2}\sum_{k=1}^{N}(u_k - R^{-1}y_k^d)^{\mathrm{T}} R (u_k - R^{-1}y_k^d)\\
&\quad - \sum_{k=0}^{N-1}\left[\frac{1}{2}\bar{\zeta}_k^{\mathrm{T}} Q\bar{\zeta}_k + (\zeta_k - \bar{\zeta}_k)^{\mathrm{T}} Q\bar{\zeta}_k\right] - \gamma,
\end{aligned}$$

where $\gamma$ is the constant defined in (10.14). Defining $\phi \triangleq -\theta$, the formulation of the dual problem $\mathcal{D}_e$ in (10.8)–(10.15) follows from (2.43)–(2.44) in Chapter 2, and the fact that $\max\theta = -\min(-\theta) = -\min\phi$ and the optimisers are the same. Also, from Theorem 2.6.4 in Chapter 2, we conclude that there is no duality gap, that is, the minimum achieved in (10.2) is equal to minus the minimum achieved in (10.8). □

We can think of (10.9)–(10.11) as the state equations of a system (running in reverse time) with input $u_k$ and output $\zeta_k$. Theorem 10.2.1 then shows that the dual of the primal estimation problem of minimisation with *constraints on the system inputs* (the process noise $w_k$) is an *unconstrained* optimisation problem using *projected outputs* $\bar{\zeta}_k$ in the objective function.

A particular case of Theorem 10.2.1 is the following result for the *unconstrained* case.

**Corollary 10.2.2** *In the case in which the variables $\hat{w}_k$ in the primal problem $\mathcal{P}_e$ are unconstrained (that is, $\Omega = R^m$), the dual problem becomes:*

$$\mathcal{D}_e: \quad \min_{\lambda_k, u_k} \frac{1}{2}\Big\{(A^{\mathrm{T}}\lambda_0 + P_0^{-1}\mu_0)^{\mathrm{T}} P_0 (A^{\mathrm{T}}\lambda_0 + P_0^{-1}\mu_0) + \sum_{k=1}^{N}(u_k - R^{-1}y_k^d)^{\mathrm{T}} R (u_k - R^{-1}y_k^d) + \sum_{k=0}^{N-1}\lambda_k^{\mathrm{T}} BQB^{\mathrm{T}}\lambda_k\Big\} + \gamma,$$

*subject to:*

$$\lambda_{k-1} = A^{\mathrm{T}}\lambda_k + C^{\mathrm{T}}u_k \quad \text{for } k = 1, \cdots, N,$$

$$\lambda_N = 0,$$

*where $\gamma$ is the constant defined in* (10.14).

*Proof.* Note that $\bar{\zeta}_k = \zeta_k$ in (10.12) since the projection (10.15) reduces to the identity mapping in the unconstrained case. The result then follows upon substituting $\bar{\zeta}_k = \zeta_k = B^{\mathrm{T}}\lambda_k$ in expression (10.13). □

## 10.3 An Equivalent Formulation of the Primal Problem

In the previous section we have shown that problem $\mathcal{D}_e$ is dual to problem $\mathcal{P}_e$ in (10.2)–(10.7). We can gain further insight by expressing $\mathcal{P}_e$ in a different way. This is facilitated by the following results.

**Lemma 10.3.1** *Let $\tilde{\Omega} \subset \mathbb{R}^m$ be a closed convex set with a nonempty interior. Let $s \in \mathbb{R}^m$ such that $s \notin \tilde{\Omega}$. Then there exists a unique point $\bar{s} \in \tilde{\Omega}$ with minimum Euclidean distance from $s$. Furthermore, $s$ and $\bar{s}$ satisfy the inequality*

$$(s - \bar{s})^{\mathrm{T}}(\bar{s} - \xi) > 0 \tag{10.39}$$

*for any point $\xi$ in the interior of $\tilde{\Omega}$.*

*Proof.* By assumption, $\tilde{\Omega}$ is a nonempty closed convex set. From Theorem 2.3.1 of Chapter 2, we have that there exists a unique $\bar{s} \in \tilde{\Omega}$ with minimum Euclidean distance from $s$, and $\bar{s}$ is the minimiser if and only if

$$(s - \bar{s})^{\mathrm{T}}(z - \bar{s}) \leq 0 \text{ for all } z \in \tilde{\Omega}. \tag{10.40}$$

Now, let $\xi \in \operatorname{int} \tilde{\Omega}$. We will show that (10.39) holds. Since $\xi \in \tilde{\Omega}$, (10.40) holds for $z = \xi$. Thus we only need to show that (10.40) for $z = \xi \in \operatorname{int} \tilde{\Omega}$ can never be an equality. Suppose, by contradiction, that

$$(s - \bar{s})^{\mathrm{T}}(\xi - \bar{s}) = 0. \tag{10.41}$$

Note that $\|s - \bar{s}\| > 0$ since $\tilde{\Omega}$ is closed, and $s \notin \tilde{\Omega}$, $\bar{s} \in \tilde{\Omega}$. Since $\xi \in \operatorname{int} \tilde{\Omega}$, there exists an $\varepsilon > 0$ such that the ball $N_\varepsilon(\xi) \triangleq \{z : \|z - \xi\| < \varepsilon\}$ is contained in $\tilde{\Omega}$. Define

$$\tilde{\xi} = \xi + \alpha \frac{s - \bar{s}}{\|s - \bar{s}\|}, \qquad 0 < \alpha < \varepsilon; \tag{10.42}$$

hence, $\|\tilde{\xi} - \xi\| = \alpha < \varepsilon$ and $\tilde{\xi} \in N_\varepsilon(\xi)$. We then have, using (10.41) and (10.42), that

$$(s - \bar{s})^{\mathrm{T}}(\tilde{\xi} - \bar{s}) = (s - \bar{s})^{\mathrm{T}}(\xi - \bar{s}) + \alpha \frac{(s - \bar{s})^{\mathrm{T}}(s - \bar{s})}{\|s - \bar{s}\|} = \alpha \|s - \bar{s}\| > 0.$$

Thus, we have found a point $\tilde{\xi} \in \tilde{\Omega}$ (since $N_\varepsilon(\xi)$ is contained in $\tilde{\Omega}$) such that $(s - \bar{s})^{\mathrm{T}}(\tilde{\xi} - \bar{s}) > 0$, which contradicts (10.40). Thus, (10.39) must be true, and the result follows. □

**Lemma 10.3.2** *Let $f : \mathbb{R}^m \to \mathbb{R}$ be any function and let $\Omega \subset \mathbb{R}^m$ be a closed convex set that contains an interior point $c$. Consider the optimisation problem*

$$\mathcal{P}_1' : \quad \min_w V(w), \tag{10.43}$$

*with*

$$V(w) \triangleq f(\bar{w}) + (w - \bar{w})^{\mathrm{T}} Q^{-1}(\bar{w} - c), \tag{10.44}$$

$$\bar{w} \triangleq Q^{1/2} \Pi_{\tilde{\Omega}} Q^{-1/2} w, \tag{10.45}$$

*where $\Pi_{\tilde{\Omega}}$ is the mapping that assigns to any vector $s$ in $\mathbb{R}^m$ the vector $\bar{s}$ in $\tilde{\Omega}$ that is closest to $s$ in Euclidean distance, that is,*

$$\begin{aligned} \Pi_{\tilde{\Omega}} : \mathbb{R}^m &\longrightarrow \tilde{\Omega} \\ s &\longmapsto \bar{s} = \Pi_{\tilde{\Omega}} s \triangleq \arg\min_{z \in \tilde{\Omega}} \|z - s\|, \end{aligned} \tag{10.46}$$

*and set $\tilde{\Omega}$ is defined as*

$$\tilde{\Omega} \triangleq \{z : Q^{1/2} z \in \Omega\}. \tag{10.47}$$

*Then,*

$$V(\bar{w}) < V(w) \quad \textit{for all } w \in \mathbb{R}^m \backslash \Omega.$$

*Proof.* Suppose that $w^* \in \mathbb{R}^m \backslash \Omega$ and let

$$\bar{w}^* \triangleq Q^{1/2}\Pi_{\tilde{\Omega}}Q^{-1/2}w^*. \tag{10.48}$$

Notice that $\bar{w}^* \in \Omega$ since (10.45), with $\Pi_{\tilde{\Omega}}$ and $\tilde{\Omega}$ defined in (10.46) and (10.47), respectively, defines a projection of $\mathbb{R}^m$ onto $\Omega$.

Define,

$$s^* \triangleq Q^{-1/2}w^*, \qquad \bar{s}^* \triangleq Q^{-1/2}\bar{w}^*. \tag{10.49}$$

Then, by construction, $s^*$ and $\bar{s}^*$ satisfy,

$$\bar{s}^* = \Pi_{\tilde{\Omega}}s^*, \tag{10.50}$$

and, in particular, $\bar{s}^* \in \tilde{\Omega}$. Using (10.48) and (10.49) in (10.44)–(10.45), we obtain,

$$V(w^*) = f(\bar{w}^*)+(w^*-\bar{w}^*)^{\mathrm{T}}Q^{-1}(\bar{w}^*-c) = f(Q^{1/2}\bar{s}^*)+(s^*-\bar{s}^*)^{\mathrm{T}}(\bar{s}^*-Q^{-1/2}c). \tag{10.51}$$

Also, since $\bar{w}^* \in \Omega$, we have $\overline{(\bar{w}^*)} \triangleq Q^{1/2}\Pi_{\tilde{\Omega}}Q^{-1/2}\bar{w}^* = \bar{w}^*$. Thus,

$$V(\bar{w}^*) = f(\bar{w}^*) + (\bar{w}^* - \bar{w}^*)^{\mathrm{T}}Q^{-1}(\bar{w}^* - c) = f(Q^{1/2}\bar{s}^*). \tag{10.52}$$

It is easy to see, from the assumptions on $\Omega$, that $\tilde{\Omega}$ in (10.47) is a closed convex set and $Q^{-1/2}c \in \operatorname{int}\tilde{\Omega}$ since $Q^{1/2} > 0$. From Lemma 10.3.1, equation (10.50), the definition of $\Pi_{\tilde{\Omega}}$ in (10.46), and noticing that $s^* \triangleq Q^{-1/2}w^* \notin \tilde{\Omega}$, we conclude that

$$(s^* - \bar{s}^*)^{\mathrm{T}}(\bar{s}^* - Q^{-1/2}c) > 0.$$

Hence, from (10.51) and (10.52), we have

$$V(w^*) - V(\bar{w}^*) = (s^* - \bar{s}^*)^{\mathrm{T}}(\bar{s}^* - Q^{-1/2}c) > 0.$$

The result then follows. □

In the sequel, we consider two optimisation problems to be equivalent if they both achieve the same optimum and if the optimisers are the same.

**Corollary 10.3.3** *Under the conditions of Lemma 10.3.2, problem $\mathcal{P}_1'$ defined by* (10.43)–(10.47) *is equivalent to the following problem*

$$\mathcal{P}_1: \quad \min_{w\in\Omega} f(w). \tag{10.53}$$

*Proof.* It follows from Lemma 10.3.2 that for any point $w$ in $\mathbb{R}^m\backslash\Omega$ we can find a point $\bar{w}$ in $\Omega$ that yields a strictly lower objective function value. Hence, we can perform the minimisation of (10.44) in $\Omega$ without losing global optimal solutions. Since the mapping $Q^{1/2}\Pi_{\tilde{\Omega}}Q^{-1/2}$ used in (10.45) reduces to the identity mapping in $\Omega$, we conclude that (10.44) is equal to the objective function in (10.53) for all $w \in \Omega$, and thus the problems are equivalent. □

**Corollary 10.3.4** *Let $f : \mathbb{R}^n \times \mathbb{R}^m \times \cdots \times \mathbb{R}^m \to \mathbb{R}$ be any function and let $\Omega \subset \mathbb{R}^m$ be a closed convex set that contains zero in its interior. Consider the optimisation problem*

$$\mathcal{P}_2' : \quad \min_{x_0, w_0, \ldots, w_{N-1}} V(x_0, w_0, \ldots, w_i, \ldots, w_{N-1}), \tag{10.54}$$

*with*

$$V(x_0, w_0, \ldots, w_i, \ldots, w_{N-1}) \triangleq f(x_0, \bar{w}_0, \ldots, \bar{w}_i, \ldots, \bar{w}_{N-1}) + \sum_{k=0}^{N-1} (w_k - \bar{w}_k)^{\mathrm{T}} Q^{-1} \bar{w}_k, \tag{10.55}$$

*and*

$$\bar{w}_i = Q^{1/2} \Pi_{\tilde{\Omega}} Q^{-1/2} w_i \quad \text{for } i = 0, \ldots, N-1, \tag{10.56}$$

*where $\Pi_{\tilde{\Omega}}$ and $\tilde{\Omega}$ are defined as in* (10.46) *and* (10.47)*, respectively.*

*Then, if $w_i \in \mathbb{R}^m \backslash \Omega$ for some $i \in \{0, \ldots, N-1\}$, we have*

$$V(x_0, w_0, \ldots, \bar{w}_i, \ldots, w_{N-1}) < V(x_0, w_0, \ldots, w_i, \ldots, w_{N-1})$$

*for all $x_0 \in \mathbb{R}^n$ and $w_0, \ldots, w_{i-1}, w_{i+1}, \ldots, w_{N-1} \in \mathbb{R}^m$.*

*Proof.* Consider the sequence $\{x_0^*, w_0^*, \ldots, w_i^*, \ldots, w_{N-1}^*\}$ and suppose $w_i^* \in \mathbb{R}^m \backslash \Omega$ for some $i$. Via a similar argument to that used in the proof of Lemma 10.3.2 (with $c = 0$), we can show that the sequence $\{x_0^*, w_0^*, \ldots, \bar{w}_i^*, \ldots, w_{N-1}^*\}$, with $\bar{w}_i^* = Q^{1/2} \Pi_{\tilde{\Omega}} Q^{-1/2} w_i^*$, gives a lower value of the objective function (10.55). The result then follows. □

**Corollary 10.3.5** *Under the conditions of Corollary 10.3.4, problem $\mathcal{P}_2'$ defined by* (10.54)–(10.56) *is equivalent to the problem*

$$\mathcal{P}_2 : \quad \min_{w_k \in \Omega, x_0} f(x_0, w_0, \ldots, w_i, \ldots, w_{N-1}). \tag{10.57}$$

*Proof.* Similar to the proof of Corollary 10.3.3. □

We are now ready to express the primal estimation problem $\mathcal{P}_e$ defined by equations (10.2)–(10.7) in an equivalent form. This is done in the following theorem.

**Theorem 10.3.6 (Equivalent Primal Formulation)** *Assume that $\Omega$ is a closed convex set that contains zero in its interior. Then the primal estimation problem $\mathcal{P}_e$ defined by equations* (10.2)–(10.7) *is equivalent to the following unconstrained optimisation problem:*

$$\mathcal{P}_e' : \quad V_N^{\text{OPT}}(\mu_0, y_k^d) \triangleq \min_{\hat{x}_k, \hat{v}_k, \hat{w}_k} V_N'(\{\hat{x}_k\}, \{\hat{v}_k\}, \{\hat{w}_k\}), \tag{10.58}$$

*subject to:*

$$\hat{x}_{k+1} = A\hat{x}_k + B\bar{w}_k \quad \text{for } k = 0, \cdots, N-1, \tag{10.59}$$

$$\hat{v}_k = y_k^d - C\hat{x}_k \quad \text{for } k = 1, \cdots, N, \tag{10.60}$$

$$\bar{w}_k = Q^{1/2} \Pi_{\tilde{\Omega}} Q^{-1/2} \hat{w}_k \quad \text{for } k = 0, \ldots, N-1, \tag{10.61}$$

*where*

$$V'_N(\{\hat{x}_k\},\{\hat{v}_k\},\{\hat{w}_k\}) \triangleq \frac{1}{2}(\hat{x}_0-\mu_0)^{\mathrm{T}}P_0^{-1}(\hat{x}_0-\mu_0)+\frac{1}{2}\sum_{k=1}^{N}\hat{v}_k^{\mathrm{T}}R^{-1}\hat{v}_k + \sum_{k=0}^{N-1}\left[\frac{1}{2}\bar{w}_k^{\mathrm{T}}Q^{-1}\bar{w}_k+(\hat{w}_k-\bar{w}_k)^{\mathrm{T}}Q^{-1}\bar{w}_k\right], \quad (10.62)$$

*where* $\Pi_{\tilde{\Omega}}$ *and* $\tilde{\Omega}$ *are defined in* (10.46) *and* (10.47)*, respectively.*

*Proof.* First note that, using the equations (10.3) and (10.4), the objective function (10.7) can be written in the form

$$V_N\big(\{\hat{x}_k\},\{\hat{v}_k\},\{\hat{w}_k\}\big) = f(\hat{x}_0,\hat{w}_0,\dots,\hat{w}_i,\dots,\hat{w}_{N-1}).$$

Since the minimisation of the above objective function is performed for $\hat{x}_0 \in \mathbb{R}^n$ and for $\hat{w}_k \in \Omega$, we conclude that problem $\mathcal{P}_e$ can be written in the form (10.57). Using Corollary 10.3.5 we can then express $\mathcal{P}_e$ in the form of problem $\mathcal{P}'_2$ defined by (10.54)–(10.56). However, this is equivalent to (10.58)–(10.62) (note the presence of $\bar{w}_k$ in (10.59)), and the result then follows. □

Theorem 10.3.6 shows that the primal estimation problem of minimisation with *constraints on the system inputs* (the process noise $w_k$) can be transformed into an equivalent *unconstrained* minimisation problem *using projected inputs* $\bar{w}_k$ both in the objective function and in the state equations (10.59).

Comparing the primal problem in its equivalent formulation (10.58)–(10.62) with the dual problem (10.8)–(10.13) we observe an interesting *symmetry* between them. This is discussed in the following section.

## 10.4 Symmetry of Constrained Estimation and Control

In summary, we have shown that the two following problems are dual in the Lagrangian sense.

**Primal Constrained Problem (Equivalent Unconstrained Form)**

$$\mathcal{P}'_e: \quad \min_{\hat{x}_k,\hat{v}_k,\hat{w}_k}\left\{\frac{1}{2}(\hat{x}_0-\mu_0)^{\mathrm{T}}P_0^{-1}(\hat{x}_0-\mu_0)+\frac{1}{2}\sum_{k=1}^{N}\hat{v}_k^{\mathrm{T}}R^{-1}\hat{v}_k + \sum_{k=0}^{N-1}\left[\frac{1}{2}\bar{w}_k^{\mathrm{T}}Q^{-1}\bar{w}_k+(\hat{w}_k-\bar{w}_k)^{\mathrm{T}}Q^{-1}\bar{w}_k\right]\right\},$$

subject to:

$\hat{x}_{k+1} = A\hat{x}_k + B\bar{w}_k$ for $k=0,\cdots,N-1$,

$\hat{v}_k = y_k^d - C\hat{x}_k$ for $k=1,\cdots,N$,

$\bar{w}_k = Q^{1/2}\Pi_{\tilde{\Omega}}Q^{-1/2}\hat{w}_k$ for $k=0,\dots,N-1$.

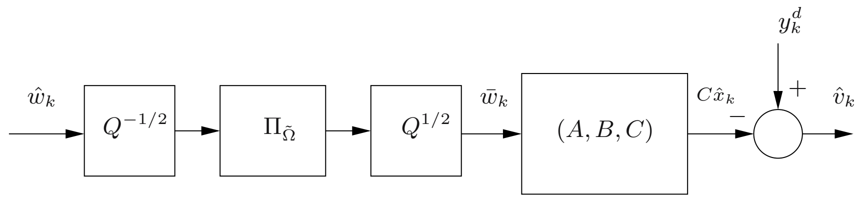


**Figure 10.1.** Configuration for the primal problem (equivalent formulation).

**Dual Unconstrained Problem**

$$\mathcal{D}_e: \quad \min_{\lambda_k, u_k} \left\{ \frac{1}{2}(\lambda_{-1} - \tilde{\mu}_0)^{\mathrm{T}} P_0 (\lambda_{-1} - \tilde{\mu}_0) + \frac{1}{2} \sum_{k=1}^{N} \hat{u}_k^{\mathrm{T}} R \hat{u}_k + \sum_{k=0}^{N-1} \left[ \frac{1}{2} \bar{\zeta}_k^{\mathrm{T}} Q \bar{\zeta}_k + (\zeta_k - \bar{\zeta}_k)^{\mathrm{T}} Q \bar{\zeta}_k \right] \right\} + \gamma,$$

subject to:

$$\lambda_{k-1} = A^{\mathrm{T}} \lambda_k + C^{\mathrm{T}} u_k \quad \text{for } k = 1, \cdots, N,$$
$$\lambda_N = 0, \quad \lambda_{-1} \triangleq A^{\mathrm{T}} \lambda_0,$$
$$\hat{u}_k \triangleq R^{-1} y_k^d - u_k \quad \text{for } k = 1, \cdots, N,$$
$$\zeta_k = B^{\mathrm{T}} \lambda_k \quad \text{for } k = 0, \cdots, N-1,$$
$$\bar{\zeta}_k = Q^{-1/2} \Pi_{\tilde{\Omega}} Q^{1/2} \zeta_k \quad \text{for } k = 0, \cdots, N-1,$$

where $\tilde{\mu}_0 \triangleq -P_0^{-1} \mu_0$ and $\gamma$ is the constant defined in (10.14).

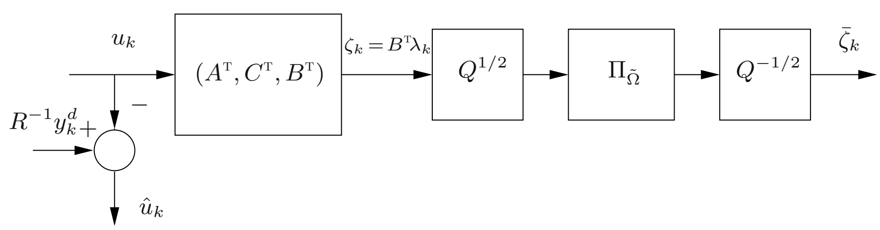


**Figure 10.2.** Configuration for the dual problem.

In the above two problems, $\Pi_{\tilde{\Omega}}$ is the minimum Euclidean distance projection defined in (10.46) onto the set $\tilde{\Omega}$ defined in (10.47).

Figures 10.1 and 10.2 illustrate the primal equivalent problem $\mathcal{P}'_e$ and the dual problem $\mathcal{D}_e$, respectively. Note from the figures and corresponding equations the symmetry between both problems; namely, input variables take the role of output variables in the objective function, system matrices are *swapped*: $A \longrightarrow A^T$, $B \longrightarrow C^T$, $C \longrightarrow B^T$, time is *reversed* and input projections become output projections.

## 10.5 Scalar Case

The above duality result takes a particularly simple form in the scalar input case, that is, when $m = 1$ in (10.1). We assume $\Omega = \{w : |w| \leq \Delta\}$, where $\Delta$ is a positive constant, and take $Q = 1$ in the objective function (10.7), without loss of generality, since we can always scale by this factor. In this case, $\tilde{\Omega} = \Omega$ and the minimum Euclidean distance projection reduces to the usual saturation function defined as $\text{sat}_\Delta(u) = \text{sign}\,(u)\,\min(|u|, \Delta)$.

The (equivalent) primal and dual problems for the scalar case are then:

**Primal Problem**

$$\begin{aligned}\mathcal{P}'_e: \quad \min_{\hat{x}_k, \hat{v}_k, \hat{w}_k} \Big\{ &\frac{1}{2}(\hat{x}_0 - \mu_0)^{\mathrm{T}} P_0^{-1} (\hat{x}_0 - \mu_0) \\ &+ \frac{1}{2}\sum_{k=1}^{N} \hat{v}_k^{\mathrm{T}} R^{-1} \hat{v}_k + \frac{1}{2}\sum_{k=0}^{N-1} \left[\hat{w}_k^2 - (\hat{w}_k - \text{sat}_\Delta(\hat{w}_k))^2\right] \Big\},\end{aligned}$$

subject to:

$$\hat{x}_{k+1} = A\hat{x}_k + B\text{sat}_\Delta(\hat{w}_k) \quad \text{for } k = 0, \cdots, N-1,$$

$$\hat{v}_k = y_k^d - C\hat{x}_k \quad \text{for } k = 1, \cdots, N.$$

**Dual Problem**

$$\begin{aligned}\mathcal{D}_e: \quad \min_{\lambda_k, u_k} \Big\{ &\frac{1}{2}(A^{\mathrm{T}}\lambda_0 + P_0^{-1}\mu_0)^{\mathrm{T}} P_0 (A^{\mathrm{T}}\lambda_0 + P_0^{-1}\mu_0) \\ &+ \frac{1}{2}\sum_{k=1}^{N} \hat{u}_k^{\mathrm{T}} R \hat{u}_k + \frac{1}{2}\sum_{k=0}^{N-1} \left[\zeta_k^2 - (\zeta_k - \text{sat}_\Delta(\zeta_k))^2\right] \Big\} + \gamma, \quad (10.63)\end{aligned}$$

subject to:

$$\lambda_{k-1} = A^{\mathrm{T}}\lambda_k + C^{\mathrm{T}}u_k \quad \text{for } k = 1, \cdots, N, \quad (10.64)$$

$$\lambda_N = 0, \quad (10.65)$$

$$\hat{u}_k \triangleq R^{-1}y_k^d - u_k \quad \text{for } k = 1, \cdots, N, \quad (10.66)$$

$$\zeta_k = B^{\mathrm{T}}\lambda_k \quad \text{for } k = 0, \cdots, N-1, \quad (10.67)$$

where $\gamma$ is the constant defined in (10.14).

**Example 10.5.1.** Consider the model (10.1) with matrices

$$A = \begin{bmatrix} 0.50 & 0.01 \\ -0.70 & 0.30 \end{bmatrix}, \ B = \begin{bmatrix} 0.40 \\ 0.90 \end{bmatrix} \text{ and } C = \begin{bmatrix} 0.90 & -0.50 \end{bmatrix}.$$

The initial state $x_0$ has a Gaussian distribution $N(\mu_0, P_0)$, with $\mu_0 = [1 \quad 2]^{\mathrm{T}}$. The output noise $\{v_k\}$ is an i.i.d. sequence having a Gaussian distribution $N(0, R)$, with $R = 0.1$. The process noise $\{w_k\}$ has a truncated Gaussian distribution of the form (9.11) in Chapter 9. For this example, we take $\Omega_1 = \Omega = \{w : |w| \leq 1\}$ and $Q = 1$. The weighting matrix $P_0$ was obtained from the steady state error covariance of the Kalman filter for the system above.

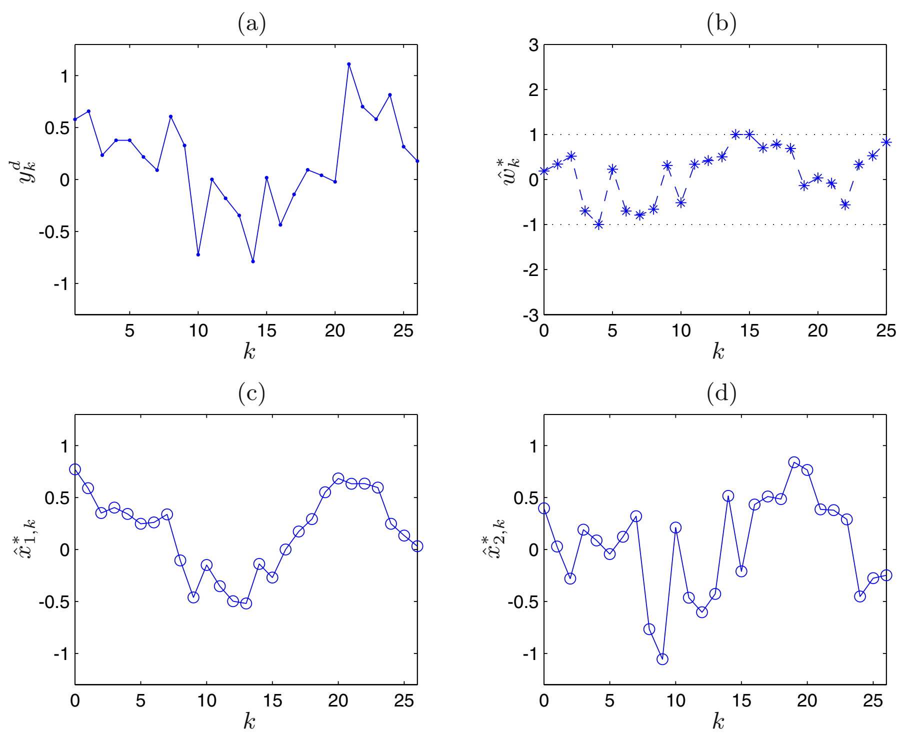


**Figure 10.3.** Primal problem: constrained estimation. (a) Measurement data. (b) Optimal primal input estimate. (c),(d) Optimal primal state estimates.

Given the measurement data $\{y_k^d\} \triangleq \{y_1^d, \cdots, y_N^d\}$ plotted in Figure 10.3 (a), we solve the primal problem (10.2)–(10.7). Note that by using equations (10.3)–(10.5) the minimisation is performed for $\hat{x}_0$ and for $\hat{w}_k \in \Omega$. We can then use QP to obtain the optimal initial state estimate $\hat{x}_0^*$, and the optimal input estimate $\hat{\omega}_k^*$. The latter is plotted in Figure 10.3 (b). In addition, using the state equations (10.3) with the optimal values $\hat{x}_0^*$ and $\hat{\omega}_k^*$, we obtain the optimal state estimates $\hat{x}^*_{1,k}$ and $\hat{x}^*_{2,k}$, $k = 0, \ldots, N$ (the two components of the state estimate vector $\hat{x}^*_k$), shown in Figure 10.3 (c) and (d), respectively.

The dual of the above estimation problem is the nonlinear optimal control problem (10.63)–(10.67), where the saturation value is $\Delta = 1$. Note that using equations (10.64)–(10.67), the decision variables of the minimisation problem (10.63) are $u_1, \ldots, u_N$ only. The dual problem has swapped the role of the inputs and outputs in the objective function. In the primal problem, the system outputs were the measurement data $y^d_k$. For the dual problem (see (10.66)), $y^d_k$ has been scaled as the input reference $R^{-1}y^d_k$ to system (10.64)–(10.65). This scaled input reference is shown in Figure 10.4 (a). We solve the nonlinear unconstrained optimisation problem (10.63)–(10.67) to obtain the optimal input $u^*_k$ shown in Figure 10.4 (b). Similarly, the dual system states $\lambda^*_k$, whose components are plotted in Figure 10.4 (c)–(d), respectively, can be obtained, in reverse time, via equations (10.64)–(10.65) by using the optimal values of $u^*_k$.

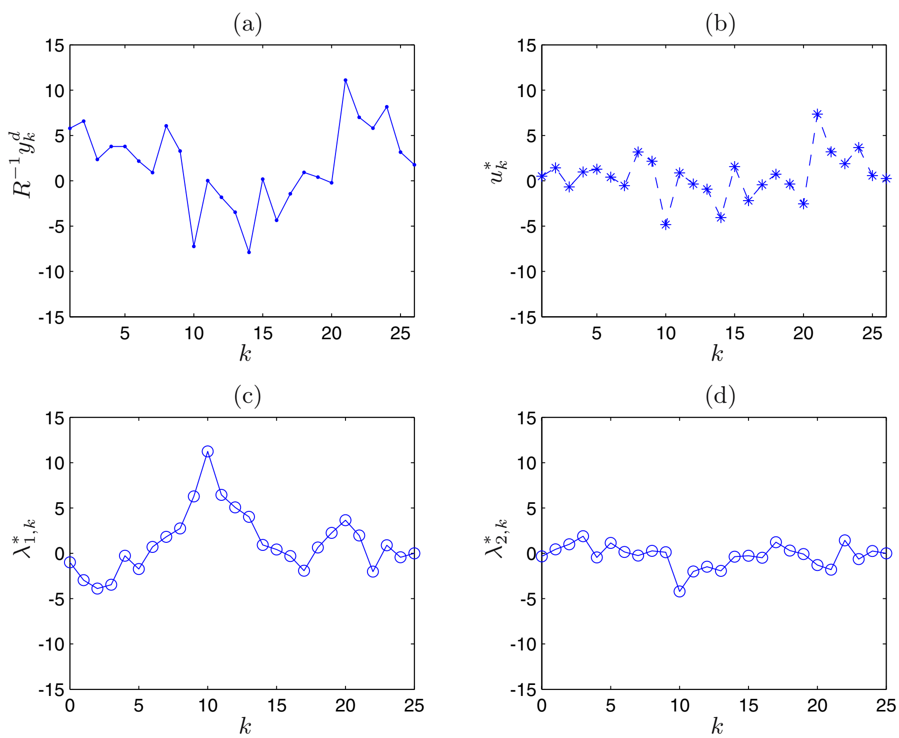


**Figure 10.4.** Dual problem: nonlinear optimal control. (a) Scaled data. (b) Optimal dual input. (c),(d) Optimal dual states.

The relation of strong Lagrangian duality between constrained estimation and control defines a relation between the optimal values in the primal and dual problems. From equation (10.67), the "dual output" $\zeta_k^* = B^{\mathrm{T}}\lambda_k^*$ (shown in Figure 10.5 (a)) is a combination of the states $\lambda_k^*$, and from the proof of Theorem 10.2.1 (see (10.31)–(10.32)), we have

$$\hat{w}_k^* = Q\bar{\zeta}_k^*, \quad \text{where} \quad \bar{\zeta}_k^* = \mathrm{sat}_\Delta(\zeta_k^*).$$

That is, the optimal input values $\hat{w}_k^*$ of the primal problem are the scaled projections of the optimal dual outputs $\zeta_k^*$, as can be seen by comparing Figure 10.3 (b) with Figure 10.5 (b). ∘

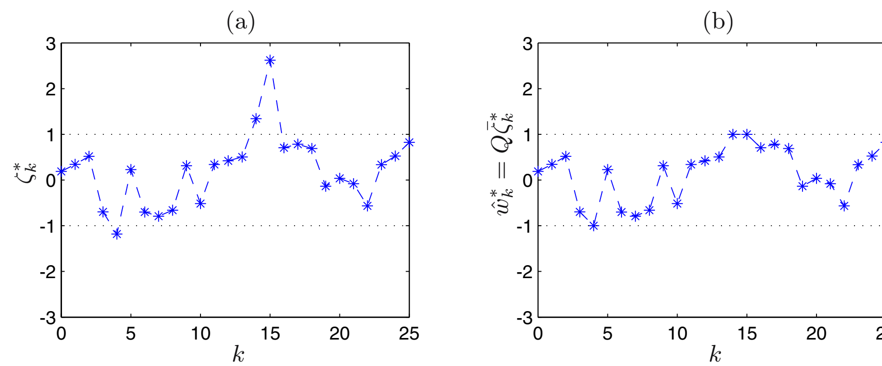


**Figure 10.5.** Relation between optimal values of the primal and dual problems. (a) Optimal dual output. (b) Optimal primal input equal to scaled projected optimal dual output.

## 10.6 More General Constraints

We have seen above that the dual of the estimation problem *with constraints on the process noise sequence* $\{w_k\}$ is an unconstrained nonlinear control problem defined in terms of projected outputs. Here we generalise the estimation problem by considering constraints on the *process noise sequence* $\{w_k\}$, the *measurement noise sequence* $\{v_k\}$ and the *initial state* $x_0$. In this case, the dual problem will turn out to be an unconstrained nonlinear control problem defined in terms of projected outputs, projected inputs and projected terminal states.

Thus, consider the following system

$$\begin{aligned} x_{k+1} &= Ax_k + Bw_k \quad \text{for } k = 0, \cdots, N-1, \\ y_k &= Cx_k + v_k \quad \text{for } k = 1, \cdots, N, \end{aligned} \tag{10.68}$$

where $x_k \in \mathbb{R}^n$, $w_k \in \mathbb{R}^m$, $y_k \in \mathbb{R}^p$. We assume that $\{w_k\}$, $\{v_k\}$, $x_0$ have truncated Gaussian distributions of the forms given in (9.11), (9.12) and (9.13) of Chapter 9, respectively, where

$$\begin{aligned} w_k &\in \Omega_1 \quad \text{for } k = 0, \cdots, N-1, \\ v_k &\in \Omega_2 \quad \text{for } k = 1, \cdots, N, \\ x_0 &\in \Omega_3. \end{aligned}$$

For (10.68) we consider the following optimisation problem:

$$\mathcal{P}_e: \qquad V_N^{\text{OPT}}(\mu_0, y_k^d) \triangleq \min_{\hat{x}_k, \hat{v}_k, \hat{w}_k} V_N(\{\hat{x}_k\}, \{\hat{v}_k\}, \{\hat{w}_k\}), \tag{10.69}$$

subject to:

$$\hat{x}_{k+1} = A\hat{x}_k + B\hat{w}_k \quad \text{for } k = 0, \ldots, N-1, \tag{10.70}$$

$$\hat{v}_k = y_k^d - C\hat{x}_k \quad \text{for } k = 1, \ldots, N, \tag{10.71}$$

$$\{\hat{x}_0, \hat{x}_1, \ldots, \hat{x}_N, \hat{v}_1, \ldots, \hat{v}_N, \hat{w}_0, \ldots, \hat{w}_{N-1}\} \in X, \tag{10.72}$$

where, in (10.72),

$$X = \Omega_3 \times \underbrace{\mathbb{R}^n \cdots \times \mathbb{R}^n}_{N} \times \underbrace{\Omega_2 \times \cdots \times \Omega_2}_{N} \times \underbrace{\Omega_1 \times \cdots \times \Omega_1}_{N}, \tag{10.73}$$

and where, in (10.69), the objective function is

$$\begin{aligned} V_N(\{\hat{x}_k\}, \{\hat{v}_k\}, \{\hat{w}_k\}) \triangleq{}& \frac{1}{2}(\hat{x}_0 - \mu_0)^{\mathrm{T}} P_0^{-1}(\hat{x}_0 - \mu_0) \\ &+ \frac{1}{2}\sum_{k=0}^{N-1} \hat{w}_k^{\mathrm{T}} Q^{-1} \hat{w}_k + \frac{1}{2}\sum_{k=1}^{N} \hat{v}_k^{\mathrm{T}} R^{-1} \hat{v}_k. \end{aligned} \tag{10.74}$$

The following result establishes duality between the constrained estimation problem $\mathcal{P}_e$ and an unconstrained nonlinear optimal control problem.

**Theorem 10.6.1 (Dual Problem)** *Assume $\Omega_1$, $\Omega_2$, $\Omega_3$ in (10.73) are nonempty closed convex sets such that there exists a feasible solution $\{\hat{x}_0, \hat{x}_1, \ldots, \hat{x}_N, \hat{v}_1, \ldots, \hat{v}_N, \hat{w}_0, \ldots, \hat{w}_{N-1}\} \in \operatorname{int} X$ for the primal problem $\mathcal{P}_e$. Given the primal constrained fixed horizon estimation problem $\mathcal{P}_e$ defined by equations (10.69)–(10.74), the Lagrangian dual problem is*

$$\mathcal{D}_e: \quad \phi^{\text{OPT}}(\mu_0, \{y_k^d\}) \triangleq \min_{\lambda_k, u_k} \phi(\{\lambda_k\}, \{u_k\}), \tag{10.75}$$

*subject to:*

$$\lambda_{k-1} = A^{\mathrm{T}}\lambda_k + C^{\mathrm{T}}u_k \quad \text{for } k = 1, \cdots, N, \tag{10.76}$$

$$\lambda_N = 0, \quad \lambda_{-1} = A^{\mathrm{T}}\lambda_0, \tag{10.77}$$

$$\zeta_k = B^{\mathrm{T}}\lambda_k \quad \text{for } k = 0, \cdots, N-1. \tag{10.78}$$

*In* (10.75)*, the objective function is*

$$\begin{aligned}\phi(\{\lambda_k\},\{u_k\}) \triangleq \frac{1}{2}(\bar{\lambda}_{-1} + P_0^{-1}\mu_0)^{\mathrm{T}} P_0(\bar{\lambda}_{-1} + P_0^{-1}\mu_0) \\ + (\lambda_{-1} - \bar{\lambda}_{-1})^{\mathrm{T}} P_0(\bar{\lambda}_{-1} + P_0^{-1}\mu_0) \\ + \sum_{k=1}^{N}\Big[\frac{1}{2}(\bar{u}_k - R^{-1}y_k^d)^{\mathrm{T}} R(\bar{u}_k - R^{-1}y_k^d) \\ + (u_k - \bar{u}_k)^{\mathrm{T}} R(\bar{u}_k - R^{-1}y_k^d)\Big] \\ + \sum_{k=0}^{N-1}\left[\frac{1}{2}\bar{\zeta}_k^{\mathrm{T}} Q\bar{\zeta}_k + (\zeta_k - \bar{\zeta}_k)^{\mathrm{T}} Q\bar{\zeta}_k\right] + \gamma \qquad (10.79)\end{aligned}$$

*where* $\gamma$ *is the constant term given by*

$$\gamma \triangleq -\frac{1}{2}\mu_0^{\mathrm{T}} P_0^{-1}\mu_0 - \frac{1}{2}\sum_{k=1}^{N}(y_k^d)^{\mathrm{T}} R^{-1} y_k^d. \qquad (10.80)$$

*In* (10.79) *the projected variables are defined as*

$$\bar{\lambda}_{-1} \triangleq P_0^{-1/2}\Pi_{\tilde{\Omega}_3} P_0^{1/2}\lambda_{-1}, \qquad (10.81)$$

$$\bar{u}_k \triangleq R^{-1/2}\Pi_{\tilde{\Omega}_2} R^{1/2} u_k \quad \textit{for } k = 1,\dots,N, \qquad (10.82)$$

$$\bar{\zeta}_k \triangleq Q^{-1/2}\Pi_{\tilde{\Omega}_1} Q^{1/2}\zeta_k \quad \textit{for } k = 0,\dots,N-1, \qquad (10.83)$$

*where* $\Pi_{\tilde{\Omega}_i}$*,* $i = 1,2,3$*, denote the minimum Euclidean distance projections (defined as in* (10.15)*) onto the sets*

$$\tilde{\Omega}_1 \triangleq \{z : Q^{1/2} z \in \Omega_1\}, \qquad (10.84)$$

$$\tilde{\Omega}_2 \triangleq \{z : R^{1/2} z \in \Omega_2\}, \qquad (10.85)$$

$$\tilde{\Omega}_3 \triangleq \{z : P_0^{1/2} z + \mu_0 \in \Omega_3\}. \qquad (10.86)$$

*Moreover, there is no duality gap, that is, the minimum achieved in* (10.69) *is equal to minus the minimum achieved in* (10.75).

*Proof.* The proof follows the same lines as the proof of Theorem 10.2.1, save that we must consider the constraints on $\hat{x}_0$ and $\hat{v}_k$, as well as on $\hat{w}_k$, when optimising (10.18). Thus, instead of (10.20) and (10.21), we need to carry out a constrained optimisation as was done for $\hat{w}_k$ in (10.19). □

## 10.7 Symmetry Revisited

We have seen in Section 10.6 that the Lagrangian dual of the general constrained estimation problem is an unconstrained nonlinear control problem involving projected variables. The symmetry in this result is revealed by transforming the primal problem into an equivalent unconstrained estimation problem using projected variables. To derive this equivalent problem, we will use the following extensions of Corollaries 10.3.4 and 10.3.5:

**Corollary 10.7.1** *Let* $f : Z \to \mathbb{R}$ *and* $h : Z \to \mathbb{R}^q$*, with* $Z = \mathbb{R}^n \times \mathbb{R}^p \times \cdots \times \mathbb{R}^p \times \mathbb{R}^m \times \cdots \times \mathbb{R}^m$*, be any functions and let* $\Omega_1 \subset \mathbb{R}^m$*,* $\Omega_2 \subset \mathbb{R}^p$ *and* $\Omega_3 \subset \mathbb{R}^n$ *be closed convex sets. Let* $0 \in \operatorname{int}\Omega_1$*,* $0 \in \operatorname{int}\Omega_2$ *and* $\mu_0 \in \operatorname{int}\Omega_3$*. Consider the optimisation problem*

$$\mathcal{P}' : \quad \min_{x_0, v_k, w_k} V(x_0, v_1, \ldots, v_N, w_0, \ldots, w_{N-1}), \tag{10.87}$$

*subject to:*

$$h(\bar{x}_0, \bar{v}_1, \ldots, \bar{v}_N, \bar{w}_0, \ldots, \bar{w}_{N-1}) = 0, \tag{10.88}$$

*with*

$$\begin{aligned} V(x_0, v_1, \ldots, v_N, w_0, \ldots, w_{N-1}) \triangleq{}& f(\bar{x}_0, \bar{v}_1, \ldots, \bar{v}_N, \bar{w}_0, \ldots, \bar{w}_{N-1}) \\ &+ (x_0 - \bar{x}_0)^{\mathrm{T}} P_0^{-1} (\bar{x}_0 - \mu_0) \\ &+ \sum_{k=1}^{N} (v_k - \bar{v}_k)^{\mathrm{T}} R^{-1} \bar{v}_k \\ &+ \sum_{k=0}^{N-1} (w_k - \bar{w}_k)^{\mathrm{T}} Q^{-1} \bar{w}_k, \end{aligned} \tag{10.89}$$

*and*

$$\bar{x}_0 = P_0^{1/2} \Pi_{\tilde{\Omega}_3} P_0^{-1/2} (x_0 - \mu_0) + \mu_0, \tag{10.90}$$

$$\bar{v}_k = R^{1/2} \Pi_{\tilde{\Omega}_2} R^{-1/2} v_k \quad \text{for } k = 1, \ldots, N, \tag{10.91}$$

$$\bar{w}_k = Q^{1/2} \Pi_{\tilde{\Omega}_1} Q^{-1/2} w_k \quad \text{for } k = 0, \ldots, N-1, \tag{10.92}$$

*where* $\Pi_{\tilde{\Omega}_i}$*,* $i = 1, 2, 3$*, are the minimum Euclidean distance projections (defined as in* (10.15)*) onto the sets* (10.84)–(10.86)*, respectively.*

*Then any global optimal solution* $\{x_0^*, v_1^*, \ldots, v_i^*, \ldots, v_N^*, w_0^*, \ldots, w_i^*, \ldots, w_{N-1}^*\}$ *of* (10.87)–(10.92) *satisfies* $x_0^* \in \Omega_3$*,* $v_i^* \in \Omega_2$ *for* $i = 1, \ldots, N$ *and* $w_i^* \in \Omega_1$ *for* $i = 0, \ldots, N-1$*.*

*Proof.* As in the proof of Corollary 10.3.4 we can show that given any feasible sequence $\{x_0^*, v_1^*, \ldots, v_i^*, \ldots, v_N^*, w_0^*, \ldots, w_i^*, \ldots, w_{N-1}^*\}$ for problem (10.87)–(10.92) and such that $x_0^* \in \mathbb{R}^n \backslash \Omega_3$, and/or $v_i^* \in \mathbb{R}^p \backslash \Omega_2$ for some $i$ and/or $w_i^* \in \mathbb{R}^m \backslash \Omega_1$ for some $i$, a lower value of the objective function is achieved by replacing these variables by their projected values $\bar{x}_0^* \in \Omega_3$, $\bar{v}_i^* \in \Omega_2$ and $\bar{w}_i^* \in \Omega_1$, computed as in (10.90)–(10.92). Since the sequence so obtained satisfies the equality constraint (10.88), it is feasible and hence the result follows. □

**Corollary 10.7.2** *Under the conditions of Corollary 10.7.1, problem $\mathcal{P}'$ defined by* (10.87)–(10.92) *is equivalent to the problem*

$$\mathcal{P}: \quad \min_{x_0 \in \Omega_3, v_k \in \Omega_2, w_k \in \Omega_1} f(x_0, v_1, \ldots, v_N, w_0, \ldots, w_{N-1}).$$

*Proof.* Similar to the proof of Corollary 10.3.3. □

We then have the following equivalent formulation for the primal estimation problem $\mathcal{P}_e$ defined by equations (10.69)–(10.74).

**Theorem 10.7.3 (Equivalent Primal Formulation Revisited)** *Suppose $\Omega_1 \subset \mathbb{R}^m$, $\Omega_2 \subset \mathbb{R}^p$ and $\Omega_3 \subset \mathbb{R}^n$ are closed convex sets such that $0 \in \operatorname{int}\Omega_1$, $0 \in \operatorname{int}\Omega_2$ and $\mu_0 \in \operatorname{int}\Omega_3$. Then the primal estimation problem $\mathcal{P}_e$ defined by equations* (10.69)–(10.74) *is equivalent to the following unconstrained optimisation problem:*

$$\mathcal{P}'_e: \quad V_N^{\text{OPT}}(\mu_0, y_k^d) \triangleq \min V'_N(\hat{x}_0, \{\hat{v}_k\}, \{\hat{w}_k\}), \tag{10.93}$$

*subject to:*

$$\bar{x}_{k+1} = A\bar{x}_k + B\bar{w}_k \quad \text{for } k = 0, \cdots, N-1, \tag{10.94}$$

$$\bar{v}_k = y_k^d - C\bar{x}_k \quad \text{for } k = 1, \cdots, N, \tag{10.95}$$

$$\bar{x}_0 = P_0^{1/2}\Pi_{\tilde{\Omega}_3} P_0^{-1/2}(\hat{x}_0 - \mu_0) + \mu_0, \tag{10.96}$$

$$\bar{v}_k = R^{1/2}\Pi_{\tilde{\Omega}_2} R^{-1/2}\hat{v}_k \quad \text{for } k = 1, \ldots, N, \tag{10.97}$$

$$\bar{w}_k = Q^{1/2}\Pi_{\tilde{\Omega}_1} Q^{-1/2}\hat{w}_k \quad \text{for } k = 0, \ldots, N-1, \tag{10.98}$$

*where $\Pi_{\tilde{\Omega}_i}$, $i = 1, 2, 3$ are the minimum Euclidean distance projections (defined as in* (10.15)*) onto the sets* (10.84)–(10.86), *respectively, and where*

$$\begin{aligned} V'_N(\hat{x}_0, \{\hat{v}_k\}, \{\hat{w}_k\}) \triangleq\ & \frac{1}{2}(\bar{x}_0 - \mu_0)^{\mathrm{T}} P_0^{-1}(\bar{x}_0 - \mu_0) \\ & + (\hat{x}_0 - \bar{x}_0)^{\mathrm{T}} P_0^{-1}(\bar{x}_0 - \mu_0) \\ & + \sum_{k=1}^{N} \left[\frac{1}{2}\bar{v}_k^{\mathrm{T}} R^{-1}\bar{v}_k + (\hat{v}_k - \bar{v}_k)^{\mathrm{T}} R^{-1}\bar{v}_k\right] \\ & + \sum_{k=0}^{N-1} \left[\frac{1}{2}\bar{w}_k^{\mathrm{T}} Q^{-1}\bar{w}_k + (\hat{w}_k - \bar{w}_k)^{\mathrm{T}} Q^{-1}\bar{w}_k\right]. \end{aligned} \tag{10.99}$$

*Proof.* Immediate from Corollary 10.7.2 on interpreting $h$ in (10.88) as

$$h(\bar{x}_0, \bar{v}_1, \ldots, \bar{v}_N, \bar{w}_0, \ldots, \bar{w}_{N-1}) = \begin{bmatrix} CA\bar{x}_0 + \bar{v}_1 + CB\bar{w}_0 - y_1^d \\ CA^2\bar{x}_0 + \bar{v}_2 + CAB\bar{w}_0 + CB\bar{w}_1 - y_2^d \\ \vdots \\ CA^N\bar{x}_0 + \bar{v}_N + \sum_{k=0}^{N-1} CA^{N-k-1}B\bar{w}_k - y_N^d \end{bmatrix}.$$

□

If we compare the equivalent form of the primal problem (10.93)–(10.99) with the dual problem (10.75)–(10.83) then aspects of the symmetry between these problems are revealed. In particular, we see that the following connections hold: time is reversed, system matrices are swapped ($A \longrightarrow A^T$, $B \longrightarrow C^T$, $C \longrightarrow B^T$), input projections become output projections and vice versa, and initial state projections become terminal state projections. These connections and other observations are summarised in Table 10.1.

| | Primal | Equivalent Primal | Dual |
|---|---|---|---|
| State equations | $\hat{x}_{k+1} = A\hat{x}_k + B\hat{w}_k$ | $\bar{x}_{k+1} = A\bar{x}_k + B\bar{w}_k$ | $\lambda_{k-1} = A^{\mathrm{T}}\lambda_k + C^{\mathrm{T}}u_k$, $\lambda_{-1} = A^{\mathrm{T}}\lambda_0$ |
| Output equation | $\hat{v}_k = y_k^d - C\hat{x}_k$ | $\bar{v}_k = y_k^d - C\bar{x}_k$ | $\zeta_k = B^{\mathrm{T}}\lambda_k$ |
| Input/output connection | Input constraints $\hat{w}_k \in \Omega_1$ | *Unconstrained* minimisation using the *projected input* $\bar{w}_k$ in the objective function. *Projected input* used in the state equations: $\bar{x}_{k+1} = A\bar{x}_k + B\bar{w}_k$. | *Unconstrained* minimisation using *projected output* $\bar{\zeta}_k$ in the objective function. |
| Output/input connection | Output constraints $\hat{v}_k \in \Omega_2$ | *Unconstrained* minimisation using the *projected output* $\bar{v}_k$ in the objective function. *Projected output* required to satisfy the output equation: $\bar{v}_k = y_k^d - C\bar{x}_k$. | *Unconstrained* minimisation using the *projected input* $\bar{u}_k$ in the objective function. |
| Initial/final state connection | Initial state constraints $\hat{x}_0 \in \Omega_3$ | *Unconstrained* minimisation using the *projected initial state* $\bar{x}_0$ in the objective function. *Projected initial state* used as initial state for the state equations. | *Unconstrained* minimisation using the *projected terminal state* $\bar{\lambda}_{-1}$ in the objective function. |

**Table 10.1.** Connections between the primal problem, its equivalent formulation and the dual problem.

## 10.8 Further Reading

For complete list of references cited, see References section at the end of book.

### General

The relationship between linear estimation and linear quadratic control is well-known in the *unconstrained* case. Since the original work of Kalman and others (see Kalman 1960b, Kalman and Bucy 1961), many authors have contributed to further understand this relationship. For example, Kailath, Sayed and Hassibi (2000) have explored duality in the unconstrained case using the geometrical concepts of dual bases and orthogonal complements. The connection between the two unconstrained optimisation problems using Lagrangian duality has also been established in, for example, the recent work of Rao (2000).

The results in the current chapter are based on Goodwin, De Doná, Seron and Zhuo (2004).

# Part II

# Further Developments