The Conditional State Distribution as a Sufficient Statistic

There are many different functions that can serve as sufficient statistics. The identity function \( S_k(I_k) = I_k \) is certainly one of them. To obtain another important sufficient statistic, we assume that the probability distribution of the observation disturbance \( v_{k+1} \) depends explicitly only on the immediately preceding state, control, and system disturbance \( x_k, u_k, w_k, \) and not on \( x_{k-1}, \ldots, x_0, u_{k-1}, \ldots, u_0, w_{k-1}, \ldots, w_0, v_{k-1}, \ldots, v_0. \) Under this assumption, it turns out that a sufficient statistic is given by the conditional probability distribution \( P_{x_k|I_k} \) of the state \( x_k, \) given the information vector \( I_k. \) In particular, we will show that for all \( k \) and \( I_k, \) we have

\[
J_k(I_k) = \min_{u_k \in U_k} H_k(P_{x_k|I_k}, u_k) = J_k(P_{x_k|I_k}),
\]

where \( H_k \) and \( J_k \) are appropriate functions.
To this end, we note an important fact that relates to state estimation of discrete-time stochastic systems: the conditional distribution $P_{x_k|I_k}$ can be generated recursively. In particular, it turns out that we can write for all $k$

$$P_{x_{k+1}|I_{k+1}} = \Phi_k(P_{x_k|I_k}, u_k, z_{k+1}),$$

where $\Phi_k$ is some function that can be determined from the data of the problem. Let us postpone a justification of this for the moment, and accept it for the purpose of the following discussion.

We note that to perform the minimization in Eq. (5.32), it is sufficient to know the distribution $P_{x_{N-1}|I_{N-1}}$ together with the distribution $P_{w_{N-1}|x_{N-1}, u_{N-1}}$, which is part of the problem data. Thus, the minimization in the right-hand side of Eq. (5.32) is of the form

$$J_{N-1}(I_{N-1}) = \min_{u_{N-1} \in U_{N-1}} H_{N-1}(P_{x_{N-1}|I_{N-1}, u_{N-1}}) = J_{N-1}(P_{x_{N-1}|I_{N-1}}),$$

for appropriate functions $H_{N-1}$ and $J_{N-1}$.

We now use induction, i.e., we assume that

$$J_{k+1}(I_{k+1}) = \min_{u_{k+1} \in U_{k+1}} H_{k+1}(P_{x_{k+1}|I_{k+1}, u_{k+1}}) = J_{k+1}(P_{x_{k+1}|I_{k+1}}),$$

for appropriate functions $H_{k+1}$ and $J_{k+1}$, and we show that

$$J_k(I_k) = \min_{u_k \in U_k} H_k(P_{x_k|I_k, u_k}) = J_k(P_{x_k|I_k}),$$

or appropriate functions $H_k$ and $J_k$.

Indeed, for a given $I_k$, the expression

$$\min_{u_k \in U_k} E \left\{ g_k(x_k, u_k, w_k) + J_{k+1}(I_{k+1}) \mid I_k, u_k \right\}$$

in the DP equation (5.33) is written as

$$\min_{u_k \in U_k} E \left\{ g_k(x_k, u_k, w_k) + J_{k+1}(\Phi_k(P_{x_k|I_k, u_k, z_{k+1}})) \mid I_k, u_k \right\}.$$
By using Bayes’ rule, this distribution can be expressed in terms of \( P_{x_k | I_k} \),
the given distributions
\[
P(w_k | x_k, u_k), \quad P(u_{k+1} | f_k(x_k, u_k, w_k), u_k, w_k),
\]
and the system equation \( x_{k+1} = f_k(x_k, u_k, w_k) \). Therefore the expression minimized over \( u_k \) can be written as a function of \( P_{x_k | I_k} \) and \( u_k \), and the DP equation (5.33) can be written as
\[
J_k(I_k) = \min_{u_k \in U_k} H_k(P_{x_k | I_k}, u_k)
\]
for a suitable function \( H_k \). Thus the induction is complete and it follows
that the distribution \( P_{x_k | I_k} \) is a sufficient statistic.

Note that if the conditional distribution \( P_{x_k | I_k} \) is uniquely determined
by another expression \( S_k(I_k) \), that is, \( P_{x_k | I_k} = G_k(S_k(I_k)) \) for an appropriate function \( G_k \), then \( S_k(I_k) \) is also a sufficient statistic. Thus, for example, if we can show that \( P_{x_k | I_k} \) is a Gaussian distribution, then the mean and the covariance matrix corresponding to \( P_{x_k | I_k} \) form a sufficient statistic.

Regardless of its computational value, the representation of the optimal policy as a sequence of functions of the conditional probability distribution \( P_{x_k | I_k} \),
\[
\mu_k(I_k) = \pi_k(P_{x_k | I_k}), \quad k = 0, 1, \ldots, N - 1,
\]
is conceptually very useful. It provides a decomposition of the optimal controller in two parts:

(a) An estimator, which uses at time \( k \) the measurement \( z_k \) and the control \( u_{k-1} \) to generate the probability distribution \( P_{x_k | I_k} \).

(b) An actuator, which generates a control input to the system as a function of the probability distribution \( P_{x_k | I_k} \) (Fig. 5.4.1).

This interpretation has formed the basis for various suboptimal control schemes that separate the controller a priori into an estimator and an actuator and attempt to design each part in a manner that seems “reasonable.” Schemes of this type will be discussed in Chapter 6.

**Alternative Perfect State Information Reduction**

By using the sufficient statistic \( P_{x_k | I_k} \) we can write the DP algorithm in
an alternative form. Using Eq. (5.34), we have for \( k < N - 1 \)
\[
J_k(P_{x_k | I_k}) = \min_{u_k \in U_k} \left[ E_{x_k, u_k, z_k+1} \{ g_k(x_k, u_k, w_k) \}
+ J_{k+1}(\Phi_k(P_{x_k | I_k}, u_k, z_{k+1}) | I_k, u_k) \right].
\]  
(5.35)
In the case where $k = N - 1$, we have

$$J_{N-1}(P_{x_{N-1}|I_{N-1}}) = \min_{u_{N-1} \in U_{N-1}} \left[ E \left\{ g_N \left( f_{N-1}(x_{N-1}, u_{N-1}, w_{N-1}) \right) + g_{N-1}(x_{N-1}, u_{N-1}, w_{N-1}) \mid I_{N-1}, u_{N-1} \right\} \right].$$  

(5.36)

This DP algorithm yields the optimal cost as

$$J^* = E \{ J_0(P_{x_0|z_0}) \},$$

where $J_0$ is obtained by the last step, and the probability distribution of $z_0$ is obtained from the measurement equation $z_0 = h_0(x_0, v_0)$ and the statistics of $x_0$ and $v_0$.

By observing the form of Eq. (5.35), we note that it has the standard DP structure, except that $P_{x_{k}|I_{k}}$ plays the role of the “state.” Indeed the role of the “system” is played by the recursive estimator of $P_{x_{k}|I_{k}}$, $P_{x_{k+1}|I_{k+1}} = \Phi_k(P_{x_k|I_k}, u_k, z_{k+1})$, and this system fits the framework of the basic problem (the role of control is played by $u_k$ and the role of the disturbance is played by $z_{k+1}$). Furthermore, the controller can calculate (at least in principle) the state $P_{x_k|I_k}$ of this system at time $k$, so perfect state information prevails. Thus the alternate DP algorithm (5.34)-(5.35) may be viewed as the DP algorithm of the perfect state information problem that involves the above system, whose state is $P_{x_k|I_k}$, and an appropriately reformulated cost function. In the absence of perfect knowledge of the state, the controller can be viewed

Figure 5.4.1 Conceptual separation of the optimal controller into an estimator and an actuator.
as controlling the “probabilistic state” $P_{x_k | I_k}$ so as to minimize the expected cost-to-go conditioned on the information $I_k$ available.

p. 262 (+15) Change $y_0'K_ky_0$ to $y_0'K_0y_0$

p. 396 (+6) The expression should read

$$\overline{\tau}_i(u) = \sum_{j=1}^{a} \int_{0}^{\infty} \tau dQ_{ij}(\tau, u),$$

p. 402 (+15), (+19), (-5) Change $g(i, u)$ to $G(i, u)$

p. 402 (-5) After Eq. (7.54), add the following sentence: If there is an “instantaneous” one-stage cost $\hat{g}(i, u)$, the term $G(i, u)$ should be replaced $\hat{g}(i, u) + G(i, u)$ in this equation.

p. 408 (+2) Change (7.6) to (7.17)

p. 451 (+12) Change $C_k N_k^{-1}$ to $C_k' N_k^{-1}$

VOLUME 2

p. 50 (+17) Change “Tsitsiklis” to “Castanon”

p. 94 (-1) Change equation to

$$J_\mu(1) = -(1 - u^2)u + (1 - u^2)J_\mu(1)$$

p. 95 (+2) Change equation to

$$J_\mu(1) = -\frac{1 - u^2}{u}.$$ 

p. 181 (-2) Change “discount factor” to “discount factor with $\alpha < 2$”

p. 190 (-1) Change “no optimal policy (stationary or not).” to “no optimal stationary policy.”