Sufficient Statistics and Optimum Control of Stochastic Systems

C. Striebel

Report written by Siva K. Gorantla

I. Sufficient Statistic

The search for a sufficient statistics is a problem of data reduction. When a great deal of data is available, it must somehow be summarized in such a way that no valuable information is lost. For a control problem, it remains a question to find out if there exists an optimal control function which depends on the data only through the sufficient statistic. In order to be useful, this property must hold for a large class of loss functions. Less technically, a loss function represents the loss (cost in money or loss in utility) associated with an estimate being ”wrong” as a function of a measure of the degree of wrongness.

The author introduces two notions of a “sufficient statistic” - an “equivalent statistic” and an “informative statistic”. The former is related to an estimation problem. This problem is completely independent of the loss function. But the latter definition of information statistics refers to the loss function. The author i) discusses the two definitions, ii) provides a method for computing the equivalent statistics sequentially, iii) considers two classes of loss functions and provides an informative statistic for these two classes. At the end, the paper has sufficient information to construct an optimum control as a function of the informative statistic for the given loss function. This method is essentially that of dynamic programming.

Decision Theory Vs Stochastic Control: This paper essentially looks at the conditions when the control problem fits into the framework of decision theory. The control problem as defined in the standard manner does not fit very neatly into the decision theoretic framework. The main difference is the absence of the parameter space in the control problem. In order to turn the control problem into a standard decision theory problem, one can identify one of the random variables as a parameter and assign to it an a priori distribution so as to recover the original problem, in other words turn it back into a random variable. The initial state vector is the most likely candidate for this treatment. Though there are no theoretical objections, this procedure is not followed as it does not give any insight into the structure of the statistical control problem. To summarize, the distinguishing characteristics of the control problem in the framework of decision theory are:

- absence of a parameter space,
- the dependence of the loss function on the random variables,
- the dependence of the distribution of the random variables on the strategy, and
- the sequential nature of the problem.

In this report, we will discuss the following:

- In Section II, an equivalent statistic is defined. A method to compute it sequentially is discussed. Then the Gaussian case is solved as an example.
- In Section III, two classes of loss functions are considered. Informative statistic is provided for these two classes. The Gaussian examples are solved for both the cases. For the Gaussian examples, the author does not consider a particular expression for the loss functions and the solution is abstract. To avoid this, in this report, we consider particular loss functions and compute exact expressions.

Notes on loss functions is included along with the problem description at the start of Section II.

II. An Equivalent Statistic

A. Problem Statement

Define a general control problem. The n-Dimensional state variables x_t constitute a random time series defined by

$$\begin{aligned} & x_0(w_0) && (1) \\ x_{t+1} &= \phi_t(x_t, u_t, w_{t+1}) \qquad t = 0, \cdots, T && (2) \end{aligned}$$

where the control law is given by

$$u_t = u_t(z_0, \cdots, z_t; u_0, \cdots, u_{t-1}) \quad t = 0, \cdots, T \tag{3}$$

and the observations

$$z_t = z_t(x_t, \epsilon_t) \quad t = 0, \cdots, T \tag{4}$$

The unknown parameters are

i) w_t are the mechanization errors and are independent random variables with known distributions.
ii) ϵ_t are the observation errors and are independent with known distributions.

The critical assumptions hold: the functions ϕ_t and z_t are assumed to be known. The function $u_t(\cdot)$ take on values in the control vector space $\mathcal{U}_t$.

For convenience, we use the following notations

$$w = (w_0, \cdots, w_{T+1}, \epsilon_0, \cdots, \epsilon_T) \tag{5}$$

$$Z_t = z_0, \cdots, z_t \tag{6}$$

$$U_t = u_0, \cdots, u_t \tag{7}$$

$$\{\mathbf{u}\} = \{u_0(\cdot), \cdots, u_T(\cdot)\} \tag{8}$$

These represent the "error" vector, observations and the control vector, and the control law (strategy) respectively.

The problem description will be complete by defining the loss function. The loss function is assumed to have the form

$$L(x_{T+1}, u_0, \cdots, u_T) = L(x_{T+1}, U_T) \tag{9}$$

B. Discussion on Loss functions:

The loss functions considered in the paper are of the following forms

$$L(x_{T+1}, u_0, \cdots, u_T) \tag{10}$$

$$\sum_{t=0}^{T+1} L_t(x_t, u_0, \cdots, u_{t-1}) \tag{11}$$

The informative statistic for the class of loss function given in (10) can be extended to the class of loss functions in (11). This work corresponds to informative statistics in Section III. Both the classes of loss functions have the final value feature, in that only one time value of the state vector may appear in each term together with the control up to but not beyond that time point. An example which fits into the frame work is the limited fuel problem.

$$L_t(x_t, u_0, \cdots, u_{t-1}) = |x_t|^2 \ \text{ if } \sum_{s=0}^{t-1} |u_s| \leq c \tag{12}$$

$$\infty \quad \text{ if } \sum_{s=0}^{t-1} |u_s| > c \tag{13}$$

The theory does not apply to other loss functions, for example a generalized quadratic loss

$$\sum_t \sum_s x_t^* Q_{ts} x_s^* + \sum_t u_t^* u_t$$

for some matrix Q_{ts} where $Q_{ts} \neq 0$ when $t \neq s$. In this case, the conditional distribution for the state vector is not sufficient. It is necessary to re-estimate all previous state vectors before each control point.

C. Optimum Control

A control law $\{\hat{u}\}$ is said to be optimum for L if it satisfies

$$EL(w; \{\hat{u}\}) = \inf_{\{u\}} EL(w; \{u\}) \tag{14}$$

A loss function of the form $L(x_{T+1}, u_0, \cdots, u_T)$ can be equivalently represented as $L(w; \{u\})$ in terms of the error vector w and the control law $\{u\}$ because of the system dynamics (2)-(4).

The final target is to find a sufficient statistic of the data so that the optimum control defined in the sense of (14) depends on the data only through this sufficient statistic. Before actually finding such statistic for optimum control, we look at another notion of sufficiency - "equivalent" statistic which has nothing to do with the loss function or optimality in (14). The rest of this section focuses on proving the following statements.

- Consider the distribution of state x_t conditioned upon all the observations up to time t (Z_t) and the control law $\{u_0(\cdot), \cdots, u_t(\cdot)\}$. Not so surprisingly, this distribution depends only on the observations Z_t and the values of control actions taken up to time t U_{t-1}. This is key in proving all the results in equivalent and informative statistic.
- Any statistic $X_t(Z_t, U_{t-1})$ which is *equivalent* to this distribution is called an equivalent statistic. Thus, the distribution by itself is an equivalent statistic.
- The equivalent statistic is shown to have useful recursive properties and X_{t+1} can be computed using X_t, the value of the control u_t and the observation z_{t+1}.

D. Key relation in proving ALL the results

Define the conditional distribution of x_t (state) given Z_t (observations up to time t) by

$$P(x_t \in A | Z_t; u_0(.), \cdots, u_{t-1}(.)) \tag{15}$$

where A is a Borel set in the vector space $\mathcal{X}$.

The first result is to show that this conditional distribution (15) depends on the control law $\{u_0(\cdot), \cdots, u_{t-1}(\cdot)\}$ only through its value U_{t-1} at Z_{t-1}. To prove this, we consider the conditional distribution x_t given Z_t in (15) and provide a recursive relation for updating it and this expression **depends only on the new observation** z_{t+1} **and the value of control** $u_t(Z_t)$ **at** Z_t **only**, not the entire control function $u_t(.)$. Essentially, we want to compute the recursive expression for the state at time t given the observations up to time t, and this can be computed as follows:

We want to compute $P(x_{t+1} \in A | Z_{t+1})$ assuming that $P(x_t | Z_t)$ is known and the control law $\{u\}$ is known/fixed. For a fixed control law $\{u\}$, the following steps follow:

- Since the distribution for the mechanization noise w_{t+1} is known, the transition probabilities from x_t to x_{t+1} are known.

$$P[x_{t+1} \in A|x_t, Z_t; u_0(.), ..., u_t(.)] = P_t(A|x_t, u_t(Z_t)) \tag{16}$$

Note that this depends only on the value of $u_t(Z_t)$.

- Since the distribution for the observation noise ϵ_t is known, probability matrix for z_t given x_t is known.

$$P[z_t \in B|x_t, Z_{t-1}; u_0(.), ..., u_t(.)] = Q_t(B|x_t) \tag{17}$$

Note that this depends only on the value of x_t.

- Assume that we know the conditional distribution x_t given Z_t. Hence from transition probability of x_{t+1} given x_t given in (16), we can compute the conditional probability of x_{t+1} given Z_t by integrating out x_t. Thus $P[x_{t+1} \in A|Z_t; u_0(.), ..., u_t(.)]$ can be computed. *This depends only on the value of* $u_t(Z_t)$.
- Thus, using conditional distribution for x_{t+1} given Z_t and z_{t+1} given x_{t+1} from (17), we can compute the joint distribution of (x_{t+1}, z_{t+1}) given the observations until time t (Z_t).
- Having known the joint distribution of (x_{t+1}, z_{t+1}), we can compute the conditional distribution of x_{t+1} given Z_{t+1} by using Bayes' rule. Thus $P(x_{t+1} \in A|Z_{t+1}; u_0(.), ..., u_t(.))$ can be computed and *this depends only on the new observation* z_{t+1} *and the value of the control* $u_t(Z_t)$ *at* Z_t *only, not on the entire function* $u_t(\cdot)$. Instead of giving the exact expression, for simplicity, we can represent the recursive expression as

$$P(x_{t+1} \in A|Z_{t+1}; u_0(.), ..., u_t(.)) = g(z_{t+1}, u_t(Z_t), P(x_t|Z_t; u_0(.), ..., u_{t-1}(.))) \tag{18}$$

This recursive relation has implications apart from the control problem. For a stochastic system of the type considered here with the controls either fixed or the control functions given, the problem of estimating the state vector from the available data is solved by the computation of the conditional distribution of x_t given Z_t and u_{t-1}. The recursive relation mentioned above provides a method of computing this distribution.

E. Main Results:

Theorem 1: There exists functions $G_t(A|Z_t, U_{t-1}, t = 0, \cdots, T$ such that

$$P[x_t \in A|Z_t, u_0(.), \cdots, u_{t-1}(.)0] = G_t(A|Z_t, U_{t-1}(Z_{t-1}))a.sZ_t \tag{19}$$

for all control functions $\{u_0(.), \cdots, u_{t-1}(.)\}$.

Proof: The proof follows from induction and using the recursive expression result (18). At time 0, the conditional distribution of x_0 given z_0 that is $G_0(A|Z_0)$ is known since it can be computed using only distribution of observation noise $Q_0(B|x_0)$ given in (17). Thus Theorem 1 holds for t=0. Assuming it holds at time t, then $G_{t+1}(A|Z_{t+1}, U_t)$ is defined constructively from $G_t(A|Z_t, U_{t-1}), z_{t+1}$ and u_t by the recursive relation given by (18). ■

A statistic $X_t(Z_t, U_{t-1})$, a function of the data available at time t, will be called *equivalent* to the distribution $G_t(A|Z_t, U_{t-1})$ if the distribution depends on the data only through X_t, that is

$$G_t(A|Z_t, U_{t-1}) = G_t(A|X_t(Z_t, U_{t-1})) \tag{20}$$

and if, in addition, $X_t(Z_t, U_{t-1})$ can be recoverd from the knowledge of the distribution $G_t(\cdot|Z_t, U_{t-1})$.

It is always an advantage to find as simple an equivalent statistic X_t as possible. If none can be found, the distribution $G_t(A)$ itself can always be used, so that in this sense an equivalent statistic X_t always exists.

a) Example: For the Gaussian case, (solved later) the conditional distribution $G_t(A|Z_t, U_{t-1})$ is normal with mean $\hat{x}_t$ and covariance K_t. The covariance does not depend on the data, so $\hat{x}_t(Z_t, U_{t-1})$, the mean is equivalent to the distribution G_t.

$$G_t(A) = \int_{x_t \in A} \frac{1}{(2\pi)^{n/2}|K_t|^{1/2}} exp\left[-\frac{1}{2}(x_t - \hat{x}_t)^* K_t (x_t - \hat{x}_t)\right] dx_t$$

and, the mean can be recovered by $\hat{x}_t = \int xG_t(dx)$.

Theorem 2: If $X_t(Z_t, U_{t-1})$ is equivalent to the conditional distribution $G_t(A|Z_t, U_{t-1})$, then X_t satisfies a recursive relation

$$X_{t+1} = \Phi_t(X_t, u_t, z_{t+1}) \tag{21}$$

Proof: The distribution $G_{t+1}(A|Z_{t+1}, U_t)$ can be found from the recursive relation for the conditional distributions (18) where $G_t(A|Z_t, U_{t-1}) = G_t(A|X_t)$. Thus $G_{t+1}(A)$ is a function of X_t, u_t and z_{t+1}. From the assumption of equivalence X_{t+1} can be found from $G_{t+1}(\cdot)$. ■

b) Example: For the Gaussian case, this recursive equation is given by the Kalman filter equations. For this case, the dynamics of the system are linear

$$x_{t+1} = \phi_t x_t + u_t + w_{t+1} \tag{22}$$

$$z_t = H_t x_t + \epsilon_t \tag{23}$$

where w_t and ϵ_t are independent Gaussian vectors of mean zero and covariance matrices C_t and R_t. Thus $p_t(x_{t+1}|x_t, u_t) = p_{w_t}(x_{t+1} - \phi_t x_t - u_t)$ and $q_t(z_t|x_t) = p_{\epsilon_t}(z_t - H_t x_t)$ are gaussian distributions of mean zero and covariance matrices C_t and R_t. Using these, the recursive relation G_t can be computed which is a gaussian density. Essentially, the mean $\hat{x}_t$ and covariance K_t can be computed. The mean can be recursively computed using the well-known Kalman filter equations.

$$\hat{x}_{t+1} = \phi_t x_t + u_t + K_{t+1} H^*_{t+1} R^{-1}_{t+1} (z_{t+1} - H_{t+1} u_t - H_{t+1} \phi_t \hat{x}_t) \tag{24}$$

$$K^{-1}_{t+1} = H^*_{t+1} R^{-1}_{t+1} H_{t+1} + (C_t + \phi_t K_t \phi^*_t)^{-1} \tag{25}$$

The equivalent statistic is $\hat{x}_t$. The matrix K_t is independent of the data Z_t and the control U_{t-1} so it can be treated as a known constant.

Additionally, this formulation includes the case of the nonlinear control, but it does not include the important case in which the mechanization error w_t depends on the control.

To summarize this section, $X_t(Z_t, U_{t-1})$ which is the conditional distribution of x_t given Z_t and U_{t-1} is an equivalent statistic and can be recursively computed. For certain cases, an equivalent statistic of lower dimension can be found. For example, in the gaussian case, the mean of the conditional distribution is an equivalent statistic.

III. Informative Statistics

Recall the loss functions of the form $L(x_{T+1}, U_T)$ and the definition of optimum control $\{\hat{u}\}$ as in (14). Y_t is said to be informative for the control problem provided there exists a control law $\{\hat{u}_0(Y_0), \cdots, \hat{u}_T(Y_T)\}$ which is optimum in the sense defined by (14) and depends on the data only through Y_t. The definition of an informative statistic Y_t can be formally given as follows:

A statistic $Y_t(Z_t, U_{t-1})$ will be called *informative* with respect to the control problem defined by L provided if there exist functionals $L^*_t(Y_t, u_t; \overline{u}_{t+1}(.), ..., \overline{u}_T(.))$ and functions $K^*_t(Z_t, U_{t-1})$ such that

$$E[L|Z_t; u_0(.), \cdots, u_t(.), \overline{u}_{t+1}(.), \overline{u}_T(.)] = K^*_t(Z_t, U_{t-1}) + L^*_t[Y_t(Z_t, U_{t-1}), u_t; \overline{u}_{t+1}(.), \overline{u}_T(.)] \; t = 0, .., T \tag{26}$$

This equation must hold for all $u_0(.), \cdots, u_t(.)$ and $\overline{u}_{t+1}(.), \overline{u}_T(.)$ which can be written as functions of Y_s, i.e.,

$$\overline{u}_s(Z_s, U_{s-1}) = \overline{u}_s(Y_s(Z_s, U_{s-1})) \quad ; s = t+1, \cdots, T$$

If Y_t is informative in the sense defined, then an optimum control can be defined recursively starting at

$t = T$ and at each step defining

$$\hat{u}_t(Y_t) = u_t^* \tag{27}$$

a value of u_t that minimized $L_t^*(Y_t, u_t; \overline{u}_{t+1}(.), \overline{u}_T(.))$.

The following results are outlined in this section.

For loss functions of the form $L(x_{T+1}, U_T)$, the statistic $Y_t = (X_t, U_{t-1})$ is informative.
For loss functions of the form $L(x_{T+1}, U_T) + \sum_{\tau=0}^{T} L_\tau(u_\tau)$, the statistic $Y_t = X_t$ is informative.

Theorem 3: For loss functions of the from

$$L(x_{T+1}, U_T)$$

the statistic $Y_t = (X_t, U_{t-1})$ is informative. Here, X_t is the equivalent statistic.

Proof: In order to prove that Y_t is informative, from definition given in (26), we need to show that there exist functions L_t^* and K_t^* such that equation (26) holds. For loss functions of the form $L(x_{T+1}, U_T)$, $K_t^* = 0$ and we can prove the existence of L_t^* by induction. The underlying principle in the proof is that of an equivalent statistic - distribution of state depends only on observations Z_t and values of control U_{t-1}. We will see the proof for $t = T$. For other induction steps, the same arguments hold and the recursive property of the equivalent statistic is also used.

Consider $t = T$,

$$\begin{aligned} E[L(x_{T+1}, U_T)|Z_T; u_0(.), \cdots, u_T(.)] &= E[L(\phi_t(x_T, u_T, w_{T+1}), U_T)|Z_T; u_0(.), \cdots, u_T(.)] \qquad (28) \\ &= \int\int L(\phi_t(x_T, u_T(Z_T), w_{T+1}), U_T(Z_T))G(dx_T|X_T)P(dw_{T+1}) \\ &= L^*(X_T, U_{T-1}, u_T) \qquad (29) \end{aligned}$$

The second step is from Theorem 1 since X_T is an equivalent statistic. The loss function is a function of x_{T+1} and U_T, which inturn are functions of x_T and the control u_T. We know that the distribution of x_t given the observations Z_t and the control law $\{u\}$ depends only on Z_t and the values of the control up to time t at Z_t (=$U_{t-1}(Z_{t-1})$). Thus, the expected value of the loss function given as considered above can be simplified in terms of (X_t, U_{t-1}, u_T). Thus, from (29), an informative statistic at time T would be of the form $Y_t = (X_t, U_{t-1})$.

Using the recursive property of the equivalent statistic given in Theorem 2, if there exists a functional of form L_{t+1}^*, we can use the same argument used for $t = T$ case and show that

$$L_t^* = E[L(x_{T+1}, U_T)|Z_t; u_0(.), \cdots, u_t(.), \overline{u}_{t+1}(.), ..., \overline{u}_T(.)]$$

holds and can be expressed in terms of L^*_{t+1}. The optimum control $\hat{u}^*_t$ can be computed according to (27) as a function of the information statistic recursively starting from $t = T$. ■

Theorem 4: For loss functions of the from

$$L(x_{T+1}, U_T) + \sum_{\tau=0}^{T} L_\tau(u_\tau)$$

the statistic $Y_t = X_t$ is informative. X_t is the equivalent statistic.

Proof: This theorem can be proved in the same way as Theorem 3 by induction using the properties of the equivalent statistic and its recursive nature. But in this case, the expected value of the loss function conditioned upon the observations till time t has a deterministic term $\sum_{\tau=0}^{t} L_\tau(u_\tau)$ and a conditional expectation of $L(x_{T+1}, U_T) + \sum_{\tau=t+1}^{T} L_\tau(u_\tau)$ given observations Z_t. Thus, define

$$K^*_t(Z_t, U_{t-1}) = \sum_{\tau=0}^{t} L_\tau(u_\tau)$$

and

$$L^*_t = E[L(x_{T+1}, U_T) + \sum_{\tau=t+1}^{T} L_\tau(u_\tau)|Z^t]$$

By defining K^*_t and L^*_t appropriately, the proof follows by induction as in Theorem 3. Thus we can compute L^*_t and the optimum control as a function of X_t recursively. ■

c) Example: Observe that we did not actually compute the optimum control in Theorems above, but given the exact expressions for loss functions, we have all the information needed to compute them. The author considers the gaussian example but he does the same as in the previous theorems by giving a set of expressions which can be used to compute optimum control. In this report, we take a particular expression for the loss function and deduce the optimum control in terms of informative statistic for two steps (t=T,T-1).

Assume the loss function of the form $L(x_{T+1}, U_T) = x_{T+1}$. It can be shown that $Y_t = (\hat{x}_t, U_{t-1})$ is an informative statistic. At $t = T$, $E[L|Z_T, u_0, ...u_T] = E[\phi_T \hat{x}_T + u_T + \eta] = \phi_T \hat{x}_T + u_T = L^*_T$. Thus L^*_T is a function of $\hat{x}_T$ and the control at $t = T$, u_T.

We need to compute L^*_t recursively. We do it for one more step $t = T - 1$.

$$\begin{aligned} E[L|Z_{T-1}; u_0(.), ..., u_{T-1}(.), \overline{u}_T(.)] &= E[E[L|Z_{T-1}, z_T; u_0(.), ..., u_{T-1}(.), \overline{u}_T(.)]] && (30)\\ &= E[L^*_T(\phi_{T-1}\hat{x}_{T-1} + u_{T-1} + \eta, U_t, \overline{u}_T)] && (31)\\ &= E[\phi_T(\phi_{T-1}\hat{x}_{T-1} + u_{T-1} + \eta) + \overline{u}_T] && (32)\\ &= \phi_T(\phi_{T-1}\hat{x}_{T-1} + u_{T-1}) + \overline{u}_T = L^*_{T-1} && (33) \end{aligned}$$

Thus L^*_t can be computed recursively. After computing L^*_t, the optimal control law $\hat{u}_t$ which minimizes

L_t^* is computed recursively starting from $t = T$ in terms of Y_t. As seen above optimal $\hat{u}_T = -\phi_T \hat{x}_T$ and $\hat{u}_{T-1} = \frac{\hat{u}_T - \phi_T \phi_{T-1} \hat{x}_{T-1}}{\phi_T}$ and so on...

d) Example: For the Gaussian example, assume the loss function of the form $L = x_{T+1} + \sum_{s=0}^{T} u_s$. It can be shown that $Y_t = \hat{x}_t$ is an informative statistic.

The proof is analogous to that of Theorem 4, in that we need to find appropriate K_t^* and L_t^* and then find an optimum control which minimizes the L_t^* recursively. In this case,

$$K_t^* = \sum_{s=0}^{t} u_s$$

$$L_T^* = L_T(u_T) + \int L_{T+1}(\phi_T \hat{x}_T + u_T + \eta) f_{T+1}(\eta) d\eta \tag{34}$$

$$= u_T + \phi_T \hat{x}_T + u_T = 2u_T + \phi_T \hat{x}_T \tag{35}$$

and computing recursively,

$$L_t^* = L_t(u_t) + \int L_{t+1}^*(\phi_t \hat{x}_t + u_t + \eta, \overline{u}_{t+1}, ..\overline{u}_T) f_{t+1}(\eta) d\eta \tag{36}$$

Thus for $t = T - 1$,

$$L_{T-1}^* = L_{T-1}(u_{T-1}) + \int L_T^*(\phi_{T-1} \hat{x}_{T-1} + u_{T-1} + \eta, \overline{u}_T) f_T(\eta) d\eta \tag{37}$$

$$= u_{T-1} + 2\overline{u}_T + \phi_T(\phi_{T-1} \hat{x}_{T-1} + u_{T-1}) \tag{38}$$

Finding the optimal control recursively, $\hat{u}_T = -\frac{1}{2}\phi_T \hat{x}_T$ using (35), $\hat{u}_{T-1} = -\frac{1}{1+\phi_T}(\phi_T \phi_{T-1} \hat{x}_{T-1} - 2\hat{u}_T)$ using (38).