EE223: Stochastic Systems: Estimation and Control. **Spring 2007**

Lecture 12 — 22nd February

Lecturer: Venkat Anantharam *scribe: Rahul Tandra*

12.1 Introduction

We consider the traditional *sequential hypothesis testing* problem and discuss it within a partially observed stochastic control framework. A decision maker takes observations from a finite set $\mathcal{Y}$. The distribution of the observations depend on the underlying hypotheses. For simplicity we discuss the binary hypothesis problem, i.e., there are two possibilities for the underlying hypotheses. Let H_0 and H_1 denote the two hypotheses. Also let $p_0(y)$ and $p_1(y)$ be the probability distribution of the observation under H_0 and H_1 respectively. In this problem, we assume that the observations are independent conditional on the underlying hypothesis. We work with a Bayesian framework and assume that the underlying hypothesis is H_0 with probability α and is H_1 with probability $1 - \alpha$, $0 \leq \alpha \leq 1$.

We assume that the decision maker can take upto M observations sequentially at cost C per observation. However, at any point he can make a decision on the true hypothesis without making further observations. Let L_1 denote the cost of deciding H_1 when the true hypothesis is H_0 and L_0 denote the cost of deciding H_0 when the true hypothesis is H_1. The goal is to design a strategy for the decision maker that minimizes the total expected cost.

12.2 Partially observed stochastic control formulation

We now formulate the above problem as a finite horizon partially observed stochastic control problem. Let $N = M + 1$ be the time horizon. The state space can be modeled as

$$\mathcal{X} = \{\Delta, \mathbf{0}, \mathbf{1}\}$$

with the following interpretation:

- Δ: represents the state when a decision has already been made in the past
- $\mathbf{0}$: denotes the state where a decision has not yet been made and the true hypothesis is H_0
- $\mathbf{1}$: denotes the state where a decision has not yet been made and the true hypothesis is H_1

Under this state space model the initial distribution is given by $[0, \alpha, 1 - \alpha]$.
Let the set of controls be denoted by $\mathcal{U} = \{c, s_0, s_1\}$ where

- c stands for the control action where the controller decides to continue and take another observation,
- s_0 denotes the control action where the controller decides on hypothesis H_0,
- s_1 denotes the control action where the controller decides on hypothesis H_1.

It is easy to see that the transition probability matrices are given as

$$\mathbb{P}(c) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \mathbb{P}(s_0) = \mathbb{P}(s_1) = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Note that these transition probability matrices will determine the state update functions $f_k(x_k, u_k, w_k)$, $k = 0, 1, \ldots, N-1$, in our general setup.

Recall that in the general setup we had the function $h_0(x_0, v_0)$ and the functions $h_k(x_k, u_{k-1}, v_k)$, $k = 1, 2, \cdots, N$. In the sequential hypothesis testing problem there is no observation at times 0 and N, and we can determine $h_k(x_k, u_{k-1}, v_k)$ for $k = 1, 2, \ldots, N-1$ using the functions $q_k(y_k|x_k, u_{k-1})$ given by

$$\begin{aligned} q_k(*|\Delta, u) &= 1 \text{ for all } u \\ q_k(y|\mathbf{0}, c) &= p_0(y), \quad y \neq * \\ q_k(y|\mathbf{1}, c) &= p_1(y), \quad y \neq * \end{aligned}$$

for $k = 1, 2, \cdots, N-1$. Here we have augmented the state space $\mathcal{Y}$ by including the symbol '$*$' which represents the case when a decision has already been made in the past and hence no further observations are taken. Note that the remaining terms of the function are zero, i.e., $q_k(y|\Delta, u) = 0$ for all $y \neq *$ and all u, and $q_k(*|\mathbf{0}, c) = q_k(*|\mathbf{1}, c) = 0$ for all $y \neq *$. We cannot have $(x_k, u_{k-1}) \in \{(0, s_0), (0, s_1), (1, s_0), (1, s_1)\}$, so it is not necessary to define $q(y_k|x_k, u_{k-1})$ for such pairs.

We now determine the cost functions $g_k(x_k, u_k)$.

$$\begin{aligned} g_k(\Delta, u) &= 0, \quad k = 0, 1, \ldots, N-1, \quad \text{for all } \ u \in \mathcal{U} \\ g_k(\mathbf{1}, c) &= C \ , \ k = 0, 1, \ldots, N-2, \\ g_k(\mathbf{0}, c) &= C \ , \ k = 0, 1, \ldots, N-2, \\ g_{N-1}(\mathbf{1}, c) &= \infty \\ g_{N-1}(\mathbf{0}, c) &= \infty \\ g_k(\mathbf{1}, s_0) &= L_0 \ , \ k = 0, 1, \ldots, N-1, \\ g_k(\mathbf{1}, s_1) &= 0 \ , \ k = 0, 1, \ldots, N-1, \\ g_k(\mathbf{0}, s_1) &= L_1 \ , \ k = 0, 1, \ldots, N-1, \\ g_k(\mathbf{0}, s_0) &= 0 \ , \ k = 0, 1, \ldots, N-1 \ , \end{aligned}$$

Notice that the control action in state $\mathbf{0}$ or state $\mathbf{1}$ at time $N-1$ has to be s_0 or s_1 only, and this has been ensured by setting $g_{N-1}(\mathbf{1}, c) = g_{N-1}(\mathbf{0}, c) = \infty$.

12.3 Optimal strategy using DP algorithm

Now we are ready to solve this problem under our general DP framework.

Remark 12.1. *At any time under any strategy Ψ, the conditional distribution of the state X_k^Ψ given $I_k^\Psi = (Y_0^\Psi, Y_1^\Psi, \cdots, Y_k^\Psi, U_0^\Psi, \cdots, U_{k-1}^\Psi)$ is of the form $[0, \alpha_k, 1-\alpha_k]$ for some $0 \le \alpha_k \le 1$ or it is $[1, 0, 0]$. For convenience we denote the distribution $[1, 0, 0]$ by $\mathbf{F}$ and $[0, \alpha_k, 1-\alpha_k]$ is denoted by $\bar{\alpha}_k$, and sometimes, abusing notation, by α_k.*

From the above remark, it is clear that the cost to go function $J_k(.)$, $k = 0, 1, \cdots, N$ is given by $J_k(\bar{\alpha}_k)$ and $J_k(\mathbf{F})$.

Remark 12.2. *The terminal cost in this problem is zero, i.e., $g_N(x_N) = 0$. Specifically, $g_N(\Delta) = 0$. This implies that $J_N(\mathbf{F}) = 0$.*

We have now converted the problem into a fully observed stochastic control problem. Under the general setup the information state evolves as

$$[\gamma_k, \theta_k, \beta_k] \cdot \mathbb{P}(u_k) \cdot D(y_{k+1}, u_k) \propto [\gamma_{k+1}, \theta_{k+1}, \beta_{k+1}]$$

where $[\gamma_k, \theta_k, \beta_k]$ is the distribution of the state at time k, $\mathbb{P}(.)$ is the transition probability matrix and $D(y_{k+1}, u_k)$ is a diagonal matrix with entries $p(y_{k+1}|x, u_k)$.

However, in our problem $[\gamma_k, \theta_k, \beta_k]$ is either $\bar{\alpha}_k$ or $\mathbf{F}$. We now determine the information state for our problem.

- If $u_k = c$

$$\begin{aligned} \mathbf{F} &\mapsto \mathbf{F} \\ \bar{\alpha}_k &\mapsto [0, \alpha_k, 1-\alpha_k] \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} - & 0 & 0 \\ 0 & p_0(y_{k+1}) & 0 \\ 0 & 0 & p_1(y_{k+1}) \end{bmatrix} \\ &= \left[0, \frac{\alpha_k p_0(y_{k+1})}{\alpha_k p_0(y_{k+1}) + (1-\alpha_k) p_1(y_{k+1})}, 1 - \frac{\alpha_k p_0(y_{k+1})}{\alpha_k p_0(y_{k+1}) + (1-\alpha_k) p_1(y_{k+1})}\right] \end{aligned} \tag{12.1}$$

- If u_k is s_0 or s_1

$$\begin{aligned} \mathbf{F} &\mapsto \mathbf{F} \\ \bar{\alpha}_k &\mapsto \mathbf{F} \end{aligned} \tag{12.2}$$

Now, all we need to do is to solve the following DP equation

$$J_k(\lambda_k) = \min_{u_k} \mathbb{E}\left[G_k(\lambda_k, u_k) + J_{k+1}\left(T_k(\lambda_k u_k, Y_{k+1})\right)\right]$$

Remark 12.3. *λ_k stands for a realization of λ_k^Ψ, the conditional law of the state X_k^Ψ given I_k^Ψ. Recall that the only λ_k of interest are those of the form $\bar{\alpha}_k$ or $\mathbf{F}$.*

Remark 12.4. *$T_k(\lambda_k, u_k, y_{k+1})$ is the offline, strategy agnostic function that gave $p_{k+1|k+1}(x_{k+1}|\eta_{k+1})$ in terms of $p_{k|k}(x_k|\eta_k)$. In our example, this function is given in Eqn. (12.1) and Eqn. (12.2).*

Remark 12.5. *$T_k(\lambda_k, u_k, Y_{k+1})$ is a random variable. Here Y_{k+1} is distributed as the observation would be if X_k had distribution λ_k and control was u_k.*

In our example

$$\begin{aligned} J_k(\mathbf{F}) &= 0 \\ J_k(\bar{\alpha}_k) &= \min\{\alpha_k L_1, (1-\alpha_k)L_0, C + A_k(\alpha_k)\} \quad , \; k = 0, 1, \ldots, N-2 \end{aligned} \tag{12.3}$$

where

$$A_k(\alpha) := \mathbb{E}\left[J_{k+1}\left(\frac{\alpha p_0(Y)}{\alpha p_0(Y) + (1-\alpha)p_1(Y)}\right)\right]$$

here the expectation is over Y which has a distribution $(\alpha p_0(y) + (1-\alpha)p_1(y), y \in \mathcal{Y})$. Further, $J_{N-1}(F) = 0$ and the equation for $J_{N-1}(\bar{\alpha}_{N-1})$ is like equation (12.3) above, except that only the first two terms show up in the minimization (because continuing is no longer an option).

Claim 12.6. *The function $A_k(\alpha)$ satisfies the following properties:*

- *$A_k(\alpha)$ is concave over $\alpha \in [0, 1]$*
- *$A_k(0) = A_k(1) = 0$*
- *$A_k(\alpha)$ is monotonically increasing in k, i.e., $A_{k-1}(\alpha) \leq A_k(\alpha)$, for all k.*

Figure 12.3 tells us the optimal strategy. The optimal control to choose at time k is defined in terms of two thresholds $0 \leq \alpha_k^{(1)} < \alpha_k^{(2)} \leq 1$ by

$$u_k^*(\alpha_k) = \begin{cases} \text{if } \alpha_k \leq \alpha_k^{(1)} & \text{decide } s_1 \\ \text{if } \alpha_k^{(1)} < \alpha_k < \alpha_k^{(2)} & \text{decide } c \\ \text{if } \alpha_k \geq \alpha_k^{(2)} & \text{decide } s_0 \end{cases}$$

12.4 Proof of Claim 12.6

It is trivial to verify that $A_k(0) = A_k(1) = 0$ for all k. Now, we prove the monotonicity of $A_k(\alpha)$.

Note that $J_N(F) = 0$. Also,

$$J_{N-1}(\bar{\alpha}) = \min\{\alpha L_1, (1-\alpha)L_0\} \; .$$

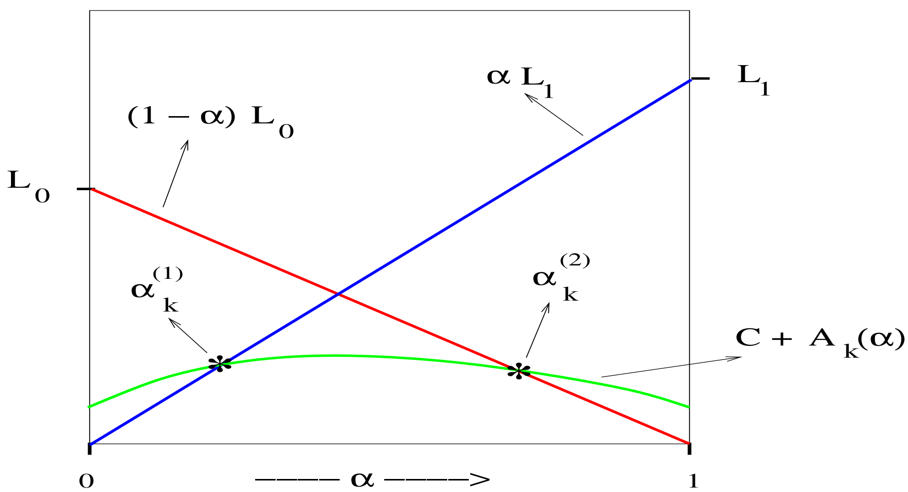

Figure 12.1. Each term in Eqn. (12.3) is plotted as a function of $\alpha \in [0, 1]$. The red and the blue straight lines correspond to the terms $(1-\alpha)L_0$ and αL_1 respectively. The green curve plots the concave function $C + A_k(\alpha)$ for some time k.

Since $J_{N-2}(\bar{\alpha})$ is a minimum of three terms, two of which are αL_1, and $(1-\alpha)L_0$, it is clear that

$$\begin{aligned} J_{N-2}(\bar{\alpha}) &\leq \min\{\alpha L_1, (1-\alpha)L_0\} \\ &= J_{N-1}(\bar{\alpha}) \end{aligned}$$

Therefore, by induction we have $J_{k-1}(\bar{\alpha}) \leq J_k(\bar{\alpha})$ for all k. From the expression for $A_k(.)$ in terms of $J_k(\cdot)$ the monotonicity of $A_k(.)$ immediately follows.

Remark 12.7. *One can use induction to show that both $J_k(\alpha)$ and $A_k(\alpha)$ are concave functions of α.*

We prove the concavity of $A_k(\alpha)$ given that of $J_{k+1}(\alpha)$. This would imply the concavity of $J_k(\alpha)$ via equation (12.3), allowing the induction to propagate. Consider $0 \leq \alpha_0, \alpha_1 \leq 1$. Define, $\alpha_\lambda := \lambda\alpha_1 + (1-\lambda)\alpha_0$.

For each $y \in \mathcal{Y}$, let

$$\begin{aligned} \xi_0(y) &:= \alpha_0 p_0(y) + (1-\alpha_0)p_1(y) \\ \xi_1(y) &:= \alpha_1 p_0(y) + (1-\alpha_1)p_1(y) \\ \xi_\lambda(y) &:= \alpha_\lambda p_0(y) + (1-\alpha_\lambda)p_1(y) \end{aligned}$$

Now, from the definition of $A_k(.)$ we have

$$\begin{aligned} A_k(\alpha_0) &= \sum_{y\in\mathcal{Y}} \xi_0(y) J_{k+1}\left(\frac{\alpha_0 p_0(y)}{\xi_0(y)}\right) \\ A_k(\alpha_1) &= \sum_{y\in\mathcal{Y}} \xi_1(y) J_{k+1}\left(\frac{\alpha_1 p_0(y)}{\xi_1(y)}\right) \\ A_k(\alpha_\lambda) &= \sum_{y\in\mathcal{Y}} \xi_\lambda(y) J_{k+1}\left(\frac{\alpha_\lambda p_0(y)}{\xi_\lambda(y)}\right) \qquad (12.4) \end{aligned}$$

Now for each $y \in \mathcal{Y}$, consider

$$\lambda\xi_1(y) \cdot J_{k+1}\left(\frac{\alpha_1 p_0(y)}{\xi_1(y)}\right) + (1-\lambda)\xi_0(y) \cdot J_{k+1}\left(\frac{\alpha_0 p_0(y)}{\xi_0(y)}\right)$$

Dividing and multiplying by $[\lambda\xi_1(y) + (1-\lambda)\xi_0(y)] =: \xi_\lambda(y)$, we get

$$\begin{aligned} & \xi_\lambda(y)\left[\frac{\lambda\xi_1(y)}{\xi_\lambda(y)} \cdot J_{k+1}\left(\frac{\alpha_1 p_0(y)}{\xi_1(y)}\right) + \frac{(1-\lambda)\xi_0(y)}{\xi_\lambda(y)} \cdot J_{k+1}\left(\frac{\alpha_0 p_0(y)}{\xi_0(y)}\right)\right] \\ \leq\ & \xi_\lambda(y)\left[J_{k+1}\left(\frac{\lambda\xi_1(y)}{\xi_\lambda(y)} \cdot \frac{\alpha_1 p_0(y)}{\xi_1(y)} + \frac{(1-\lambda)\xi_0(y)}{\xi_\lambda(y)} \cdot \frac{\alpha_0 p_0(y)}{\xi_0(y)}\right)\right] \\ =\ & \xi_\lambda(y)\left[J_{k+1}\left(\frac{[\lambda\alpha_1 + (1-\lambda)\alpha_0]p_0(y)}{\xi_\lambda(y)}\right)\right] \\ =\ & \xi_\lambda(y)\left[J_{k+1}\left(\frac{\alpha_\lambda p_0(y)}{\xi_\lambda(y)}\right)\right] \end{aligned}$$

Note, that the inequality in the second step is justified because $J_{k+1}(\bar{\alpha})$ is a concave function of α. Using the above inequality in Eqn. (12.4) will give

$$\lambda A_k(\alpha_1) + (1-\lambda)A_k(\alpha_0) \leq A_k(\alpha_\lambda)$$

which proves that $A_k(\alpha)$ is concave in α.