6.433 Recursive Estimation
6.435 System Identification

S.K. Mitter

February 13, 2002

Problem Set 1

1. Kalman Filter as an Application of the Bayes Formula

   (a) Let X be a Gaussian random variable with mean μ and variance $\sigma_x > 0$. We denote this by saying that the probability law of

   $$P(X) := N(\mu, \sigma_x) \ .$$

   Let Y be a random variable such that

   $$P(Y|X) := N(a + bX, \sigma_u) \ , \ \ \sigma_u > 0 \ .$$

   Then

   $$P(X|Y) = N(\hat{X}, \sigma_v)$$

   where $\sigma_v > 0$ and the random variable $\hat{X}$ are determined by

   $$\frac{1}{\sigma_v} = \frac{1}{\sigma_x} + \frac{b^2}{\sigma_u} \ , \ \ \frac{\hat{X}}{\sigma_v} = \frac{\mu}{\sigma_x} + \frac{b(Y-a)}{\sigma_u}$$

   (b) Show that

   $$E\{(X - \hat{X})^2\} \ = \ \sigma_v$$

   and

   $$E\{(X - \hat{X})^2 | Y\} \ = \ \sigma_v \ \text{(Conditional Expectation)}$$

   (c) Noisy Observation of a Single Random Variable
   Let $X, U_1, U_2, \ldots$ be independent random variables with

   $$P(X) = N(0, \sigma_x^2) \ , \ P(U_i) = N(0, 1) \ .$$

   Let

   $$Y_K \ = \ X + g_K U_K \ , \ \ K = 1, 2, \ldots$$
   $$g_K \ > \ 0$$

We regard the Y_K's as noisy observations of the single random variable. Let

$$\Pi_0(X) \quad = \quad P(X) := N(0, \sigma_x^2)$$

and

$$\Pi_n(X) \quad := \quad P(X|Y_1, \ldots, Y_n) \;\; n = 1, 2, \ldots$$

Prove by induction that

$$\Pi_n(X) = N(\hat{X}_n, \Sigma_n) \;\; n = 1$$

where

$$\frac{1}{\Sigma_n} = \frac{1}{\Sigma_{n-1}} + \frac{1}{g_n^2}$$

$$\frac{\hat{X}_n}{\Sigma_n} = \frac{\hat{X}_{n-1}}{\Sigma_{n-1}} + \frac{Y_n}{g_n^2}$$

(d) Now suppose that

$$X_{n+1} = f_n X_n + g_n U_n \;\; , \;\; n = 0, 1, 2, \ldots$$

$$Y_n = h_n X_n + V_n$$

Assume that $X_0, U_0, U_1, \ldots, V_0, V_1, \ldots$ are mutually independent with

$$P(X_0) \;:=\; N(0, \sigma^2)$$

$$P(U_i) \;:=\; N(0, 1)$$

and

$$P(V_i) \;:=\; N(0, 1) \;\; .$$

Let

$$\Pi_n(X_n) = P(X_n | Y_0, \ldots, Y_n) \;\; .$$

Show that $\Pi_n(X_n) := N(\hat{X}_n, \Sigma_n)$ and obtain a recursion relationship for $\hat{X}_n$ and Σ_n.

(e) Let $E_n = X_n - \hat{X}_n$. Compute the covariance of E_n, and show that it is independent of the Y's.

2. This problem is concerned with least-squares problems where there is uncertainty in the data. Let $b \in \mathbb{E}^m$ and let A be an $n \times m$ matrix which has full column rank (so $n \leq m$). We are required to solve the min-max least-squares problem

$$\operatorname*{Min}_{x} \; \operatorname*{Max}_{||\delta A||_2 \leq \gamma \, ; \, ||\delta b||_2 \leq \lambda} \left|\left| (A + \delta A)x - (b + \delta b) \right|\right|_2$$

where for a vector $b \in \mathbb{E}^m$, $|| \cdot ||_2$ is the Euclidean norm, and for a matrix A, $||A||_2$ is the maximum singular value of A.

The value of the min-max is given by

$$||A\hat{x} - b||_2 + \gamma||\hat{x}||_2 + \lambda \ ,$$

where

$$\hat{x} = (A^T A + \gamma I) A^T b \ .$$

(a) First prove that, for $x \in \mathbb{E}^n$

$$\operatorname*{Max}_{||\delta A||_2 \leq \gamma \, ; \, ||\delta b||_2 \leq \lambda} ||(A + \delta A)x - b + \delta b||d_2 = ||Ax - b||_2 + \gamma||x||_2 + \lambda \ .$$

To prove this, fix x and δA, and carry out the maximization with respect to δb satisfying the constraint. Then for fixed x and the maximizing δb, carry out the maximization with respect to δA satisfying the constraint. To do this use a singular value decomposition for $(\delta A)_0$ which performs the maximization under consideration.

(b) Finally prove that the $\hat{x}$ which carries out the minimization is given by

$$\hat{x} = (A^T A + \gamma I)^{-1} A^T b \ .$$