MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Electrical Engineering and Computer Science

6.241: Dynamic Systems—Fall 2003

Homework 2 Solutions

Exercise 2.1 The given information can be written as

$$\mathbf{y} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} r_1^2 & s_1^2 & r_1 s_1 \\ r_2^2 & s_2^2 & r_2 s_2 \\ \vdots & \vdots & \vdots \\ r_{10}^2 & s_{10}^2 & r_{10} s_{10} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{Ax}.$$

We would like to find the solution $\mathbf{x}$ that minimizes the Euclidean norm of $\mathbf{y} - \mathbf{Ax}$. The set of equations above is inconsistent and leads to the least squares solution $\mathbf{x_{LS}} = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{y}$, and can be computed in MATLAB in the 4 different ways described by parts a)-d). MATLAB often leads to numerical errors when matrices are illconditioned, and thus computing $\mathbf{x_{LS}}$ in the 4 different ways may lead to slightly different answers. However, in this case the matrix $\mathbf{A}$ is not ill-conditioned (the condition number of this matrix is 2.6931). Figure 1 plots the estimated ellipse.

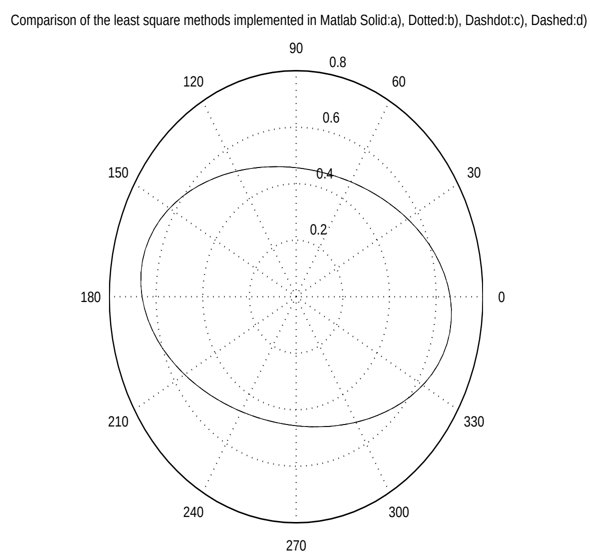

Figure 1: Figure for 2.1

The condition number of this matrix is small, 2.6931; hence, all the computational methods give us very similar solutions.

Exercise 2.2 (a) 1. For the 15th order polynomial: $f(t_i) = p_{15}(t_i) + e_i \quad i = 1, \ldots, 16$, and $t_i \in T$. Then we can express this in terms of matrix equation as follows:

$$\begin{bmatrix} y_1 \\ \vdots \\ y_{16} \end{bmatrix} = \begin{bmatrix} 1 & t_1 & t_1^2 & \cdots & t_1^{15} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_{16} & t_{16}^2 & \cdots & t_{16}^{15} \end{bmatrix} \begin{bmatrix} a_0 \\ \vdots \\ a_{15} \end{bmatrix} + \begin{bmatrix} e_1 \\ \vdots \\ e_{16} \end{bmatrix}$$

thus it can be rewritten in matrix form as $\underline{y} = A\underline{a}+\underline{e}$. But, observe that the matrix A is invertible for distinct $t_i's$. So, the coefficients a_i for $i = 0, \cdots, 15$ are $\underline{a} = A^{-1}y$, where $\underline{a} = \begin{bmatrix} a_0 & a_1 & \cdots & a_{15} \end{bmatrix}'$, and $\underline{e} = 0$.

2. For the 2nd order polynomial: $f(t_i) = p_2(t_i) + e_i \quad i = 1, \ldots, 16$, and $t_i \in T$. Then we can express the relationship between y_i and the polynomial as follows;

$$\begin{bmatrix} y_1 \\ \vdots \\ y_{16} \end{bmatrix} = \begin{bmatrix} 1 & t_1 & t_1^2 \\ \vdots & \vdots & \vdots \\ 1 & t_{16} & t_{16}^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ \vdots \\ e_{16} \end{bmatrix}$$

where the coefficients a_0, a_1, and a_2 are determined by $\underline{a} = (A'A)^{-1}A'y$, where $\underline{a} = \begin{bmatrix} a_0 & a_1 & a_2 \end{bmatrix}'$.

b) Now we have measurements affected by some noise. Then the data is

$$\tilde{y}_i = f(t_i) + n(t_i) \quad i = 1, \ldots, 16 \quad t_i \in T$$

where the noise $n(t_i)$ is generated by a command "randn" in Matlab. So, in matrix form, we have that

$$\underline{\tilde{y}} = A\underline{a} + \underline{n},$$

and a stochastic derivation shows that the "minimum variance unbiased estimator" for $\underline{a}$ is $\underline{\hat{a}} = argmin\|\underline{\tilde{y}} - A\underline{a}\|_W^2$ where $W = R_n^{-1}$, and R_n is the covariance matrix of the random variable $\underline{n}$. So,

$$\underline{\hat{a}} = (A'WA)^{-1}A'W\underline{\tilde{y}}.$$

Roughly speaking, this is saying that measurments with more noise are given less weight in the estimate of $\underline{a}$. In our problem, $R_n = I$ because the $n_i's$ are independent, zero mean and have unit variance. That is, each of the measurments is "equally noisy" or treated as equally reliable. Thus, our coefficient $\underline{\hat{a}}$ matrix is

$$\underline{\hat{a}} = (A'A)^{-1}A'\underline{\tilde{y}}$$

c) $p_2(t)$ can be written as

$$p_2(t) = a_0 + a_1 t + a_2 t^2.$$

In order to minimize the approximation error in least square sense, the optimal $\hat{p}_2(t)$ must be such that the error, $f - \hat{p}_2$, is orthogonal to every linear combination of $\{1,\ t,\ t^2\}$, i.e., that

$$< f - \hat{p}_2, 1 >= 0 \rightarrow < f, 1 >=< \hat{p}_2, 1 >$$

$$< f - \hat{p}_2, t >= 0 \rightarrow < f, t >=< \hat{p}_2, t >$$

$$< f - \hat{p}_2, t^2 >= 0 \rightarrow < f, t^2 >=< \hat{p}_2, t^2 > .$$

We have that $f = \frac{1}{2}e^{0.8t}$ for $t \in [0, 2]$, So,

$$< f, 1 >= \int_0^2 \frac{1}{2}e^{0.8t}dt = \frac{5}{8}e^{\frac{8}{5}} - \frac{5}{8}$$

$$< f, t >= \int_0^2 \frac{t}{2} e^{0.8t} dt = \frac{15}{32} e^{\frac{8}{5}} - \frac{25}{32}$$

$$< f, t^2 >= \int_0^2 \frac{t^2}{2} e^{0.8t} dt = \frac{85}{64} e^{\frac{8}{5}} - \frac{125}{64}.$$

And,

$$< \hat{p_2}, 1 >= 2a_0 + 2a_1 + \frac{8}{3} a_2$$

$$< \hat{p_2}, t >= 2a_0 + \frac{8}{3} a_1 + 4a_2$$

$$< \hat{p_2}, t^2 >= \frac{8}{3} a_0 + 4a_1 + \frac{32}{5} a_2$$

Therefore the problem reduces to solving another set of linear equations:

$$\begin{bmatrix} 2 & 2 & \frac{8}{3} \\ 2 & \frac{8}{3} & 4 \\ \frac{8}{3} & 4 & \frac{32}{5} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} < f, 1 > \\ < f, t > \\ < f, t^2 > \end{bmatrix}$$

and the solution is

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 14.5978 \\ -37.2968 \\ 17.9508 \end{bmatrix},$$

which is quite different from the solutions obtained from either a) or b). This is expected, because here we use a different notion for the closeness of the approximating polynomial, $\hat{p}_2(t)$, to the original function, f. Roughly speaking, in parts (a) and (b), the optimal polynomial will be the one for which there is smallest discrepancy between $f(t_i)$ and $p_2(t_i)$ for all t_i, i.e., the polynomial that will come closest to passing through all the sample points, $f(t_i)$. All that matters is the 16 sample points, $f(t_i)$. In this part however, all the points of f matter.

Exercise 3.1 The first and the third facts given in the problem are the keys to solve this problem, in addition to the fact that:

$$UA = \begin{pmatrix} R \\ 0 \end{pmatrix}.$$

Here note that R is a nonsingular, upper-triangular matrix so that it can be inverted. Now the problem reduces to show that

$$\hat{x} = arg \min_x \|y - Ax\|_2^2 = arg \min_x (y - Ax)'(y - Ax)$$

is indeed equal to

$$\hat{x} = R^{-1} y_1.$$

Let's transform the problem into the familiar form. We introduce an error e such that

$$y = Ax + e,$$

and we would like to minimize $\|e\|_2$ which is equivalent to minimizing $\|y - Ax\|_2$. Using the property of an orthogonal matrix, we have that

$$\|e\|_2 = \|Ue\|_2.$$

Thus with $e = y - Ax$, we have

$$\begin{aligned}\|e\|_2^2 &= \|Ue\|_2^2 = e'U'Ue = (U(y-Ax))'(U(y-Ax)) = \|Uy - UAx\|_2^2 \\ &= \left\| \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} - \begin{pmatrix} R \\ 0 \end{pmatrix} x \right\|_2^2 = (y_1 - Rx)'(y_1 - Rx) + y_2'y_2.\end{aligned}$$

Since $\|y_2\|_2^2 = y_2'y_2$ is just a constant, it does not play any role in this minimization. Thus we would lik to have

$$y_1 - R\hat{x} = 0$$

and because R is an invertible matrix, $\hat{x} = R^{-1}y_1$.

Exercise 3.2 i) We would like to minimize the 2-norm of u, i.e., $\|\underline{u}\|_2^2$. Since y_n is given as

$$y_n = \sum_{i=1}^{n} h_i u_{n-1}$$

we can rewrite this equality as

$$y_n = \begin{bmatrix} h_1 & h_2 & \cdots & h_n \end{bmatrix} \begin{bmatrix} u_{n-1} \\ u_{n-2} \\ \vdots \\ u_0 \end{bmatrix}$$

We want to find the $\underline{u}$ with the smallest 2-norm such that

$$\bar{y} = A\underline{u}.$$

where we assume that A has a full rank (i.e. $h_i \neq 0$ for some i, $1 \leq i \leq n$). Then, the solution reduces to the familiar form:

$$\hat{u} = A'(AA')^{-1}\bar{y}.$$

By noting that $AA' = \sum_{i=1}^{n} h_i^2$, we can obtain $\hat{u}_j$ as follows;

$$\hat{u}_j = \frac{h_j \bar{y}}{\sum_{i=1}^{n} h_i^2}, \quad \text{for j} = 0, 1, \cdots, \text{n} - 1.$$

ii) a) Let's introduce e as an error such that $y_n = \overline{y} - e$. It can also be written as $\overline{y} - y_n = e$. Then now the quantity we would like to minimize can be written as

$$r(\overline{y} - y_n)^2 + u_0^2 + \cdots + u_{n-1}^2$$

where r is a positive weighting parameter. The problem becomes to solve the following minimization problem :

$$\hat{u} = arg \min_u \sum_{i=1}^{n} u_i^2 + re^2 = arg \min_u (\|\underline{u}\|_2^2 + r\|e\|_2^2),$$

from which we see that r is a weight that characterizes the tradeoff between the size of the final error, $\bar{y} - y_n$, and energy of the input signal, $\underline{u}$.

In order to reduce the problem into the familiar form, i.e, $\|y - Ax\|$, let's augment $\sqrt{r}e$ at the bottom of $\underline{u}$ so that a new augmented vector, $\tilde{\underline{u}}$ is

$$\tilde{\underline{u}} = \begin{bmatrix} \underline{u} \\ \cdots \\ \sqrt{r}e \end{bmatrix},$$

This choice of $\tilde{\underline{u}}$ follows from the observation that this is the $\tilde{\underline{u}}$ that would have $\|\tilde{\underline{u}}\|_2^2 = \|\underline{u}\|_2^2 + re^2$, the quantity we aim to minimize .

Now we can write $\overline{y}$ as follows

$$\overline{y} = \begin{bmatrix} A & \vdots & \frac{1}{\sqrt{r}} \end{bmatrix} \begin{bmatrix} \underline{u} \\ \cdots \\ \sqrt{r}e \end{bmatrix} = \tilde{A}\tilde{\underline{u}} = A\underline{u} + e = y_n + e.$$

Now, $\hat{u}$ can be obtained using the augmented A, $\tilde{A}$, as

$$\hat{u} = \tilde{A}'(\tilde{A}\tilde{A}')^{-1}\overline{y} = \begin{bmatrix} A' \\ \frac{1}{\sqrt{r}} \end{bmatrix} \left[AA' + \frac{1}{r} \right] \overline{y}.$$

By noting that

$$AA' + \frac{1}{r} = \sum_{i=1}^{n} h_i^2 + \frac{1}{r},$$

we can obtain $\hat{u}_j$ as follows

$$\hat{u}_j = \frac{h_j \overline{y}}{\sum_{i=1}^{n} h_i^2 + \frac{1}{r}} \quad \text{for j} = 0, \cdots, \text{n} - 1.$$

ii) b) When $r = 0$, it can be interpreted that the error can be anything, but we would like to minimize the input energy. Thus we expect that the solution will have all the $u_i's$ to be zero. In fact, the expression obtained in ii) a) will be zero as $r \to 0$. On the other hand, the other situation is an interesting case. We put a weight of ∞ to the final state error, then the expression from ii) a) gives the same expression as in i) as $r \to \infty$.

Exercise 3.3 This problem is similar to Example 3.4, except now we require that $\dot{p}(T) = 0$. We can derive, from $x(t) = \ddot{p}(t)$, that $p(t) = x(t) * tu(t) = \int_0^t (t-\tau)x(\tau)\mathrm{d}\tau$ where $*$ denotes convolution and $u(t)$ is the unit step, defined as 1 when $t > 0$ and 0 when $t < 0$. (One way to derive this is to take $x(t) = \ddot{p}(t)$ to the Laplace domain, *taking into account initial conditions*, to find the transfer function $H(s) = P(s)/X(s)$, hence the impulse response, $h(t)$ such that $p(t) = x(t) * h(t)$). Similarly, $\dot{p}(t) = x(t) * u(t) = \int_0^t x(\tau)\mathrm{d}\tau$. So, $y = p(T) = \int_0^T (T-\tau)x(\tau)\mathrm{d}\tau$ and $0 = \dot{p}(T) = x(t) * u(t) = \int_0^T x(\tau)\mathrm{d}\tau$. You can check that $< g(t), f(t) >= \int_0^T g(t)f(t)\mathrm{d}\tau$ is an inner product on the space of continuous functions on $[0, T]$, denoted by $C[0, T]$, which we are searching for $x(t)$. So, we have that $y = p(T) =< (T-t), x(t) >$ and $0 = \dot{p}(T) =< 1, x(t) >$. In matrix form,

$$\begin{bmatrix} y \\ 0 \end{bmatrix} = \begin{bmatrix} < T-t, x(t) > \\ < 1, x(t) > \end{bmatrix} = \prec \begin{bmatrix} T-t & 1 \end{bmatrix}, x(t) \succ$$

where $\prec .,. \succ$ denotes the Grammian, as defined in chapter 2. Now, in chapter 3, it was shown that the minimum length solution to $y = \prec A, x \succ$, is $\hat{x} = A \prec A, A \succ^{-1} y$. So, for our problem,

$$\hat{x} = \begin{bmatrix} T-t & 1 \end{bmatrix} \prec \begin{bmatrix} T-t & 1 \end{bmatrix}, \begin{bmatrix} T-t & 1 \end{bmatrix} \succ^{-1} \begin{bmatrix} y \\ 0 \end{bmatrix}.$$

Where, using the definition of the Grammian, we have that:

$$\prec \begin{bmatrix} T-t & 1 \end{bmatrix}, \begin{bmatrix} T-t & 1 \end{bmatrix} \succ = \begin{bmatrix} < T-t, T-t > & < T-t, 1 > \\ < 1, T-t > & < 1, 1 > \end{bmatrix}.$$

Now, we can use the definition for inner product to find the individual entries, $< T-t, T-t >= \int_0^T (T-t)^2 \mathrm{d}t = T^3/3$, $< T-t, 1 >= \int_0^T (T-t)\mathrm{d}t = T^2/2$, and $< 1, 1 >= T$. Plugging these in, one can simplify the expression for $\hat{x}$ and obtain $\hat{x}(t) = \frac{12y}{T^2}[\frac{1}{2} - \frac{t}{T}]$ for $t \in [0, T]$.

Alternatively, we have that $x(t) = \ddot{p}(t)$. Integrating both sides and taking into account that $p(0) = 0$ and $\dot{p}(0) = 0$, we have $p(t) = \int_0^t \int_0^{t_1} x(\tau)\mathrm{d}\tau \; \mathrm{d}t_1 = \int_0^t f(t_1)\mathrm{d}t_1$. Now, we use the integration by parts formula, $\int_0^t u \; \mathrm{d}v = uv|_0^t - \int_0^t v \; \mathrm{d}u$, with $u = f(t_1) = \int_0^{t_1} x(\tau)\mathrm{d}\tau$, and $\mathrm{d}v = \mathrm{d}t_1$; hence $\mathrm{d}u = \mathrm{d}f(t_1) = x(t_1)\mathrm{d}t_1$ and $v = t_1$. Plugging in and simplifying we get that $p(t) = \int_0^t \int_0^{t_1} x(\tau)\mathrm{d}\tau \; \mathrm{d}t_1 = \int_0^t (t-\tau)x(\tau)\mathrm{d}\tau$. Thus, $y = p(T) = \int_0^T (T-\tau)x(\tau)\mathrm{d}\tau =< T-t, x(t) >$. In addition, we have that $0 = \dot{p}(T) = \int_0^T x(\tau)\mathrm{d}\tau =< 1, x(t) >$. That is, we seek to find the minimum length $x(t)$ such that

$$\begin{aligned} y &= < T-t, x(t) > \\ 0 &= < 1, x(t) > . \end{aligned}$$

Recall that the minimum length solution $\hat{x}(t)$ must be a linear combination of $T-t$ and 1, i.e., $\hat{x}(t) = a_1(T-t) + a_2$. So,

$$\begin{aligned} y &= < T-t, a_1(T-t) + a_2 > &&= a_1 \int_0^T (T-t)^2 \mathrm{d}t + a_2 \int_0^T (T-t)\mathrm{d}t &&= a_1 \frac{T^3}{3} + a_2 \frac{T^2}{2} \\ 0 &= < 1, a_1(T-t) + a_2 > &&= \int_0^T (a_1(T-t) + a_2)\mathrm{d}t &&= a_1 \frac{T^2}{2} + a_2 T. \end{aligned}$$

This is a system of two equations and two unknowns, which we can rewrite in matrix form:

$$\begin{bmatrix} y \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{T^3}{3} & \frac{T^2}{2} \\ \frac{T^2}{2} & T \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix},$$

So,

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \frac{T^3}{3} & \frac{T^2}{2} \\ \frac{T^2}{2} & T \end{bmatrix}^{-1} \begin{bmatrix} y \\ 0 \end{bmatrix}.$$