CDS 112: Winter 2014/2015

Solution #5

Problem #1: This problem seeks to build intuition about how estimators work by learning about and applying the **Least Squares Estimation** technique

We assume that $y \in \mathbb{R}$ is related linearly to variable $x \in \mathbb{R}^n$ via the relationship:

$$y = \theta^T x = \sum_{i=1}^{n} \theta_i x_i$$

Given N noisy measurements of the form (x_i, y_i), the goal is to θ. Because the measurements aren't perfect, there may not exist a θ that will satisfy

$$y_i = \theta^T x_i \qquad i = 1, 2, \ldots N$$

We wish to "guess" (estimate) the $\hat{\theta}$ that might be close to θ by minimizing the cost

$$J(\theta) = \frac{1}{2} \sum_{i=1}^{N} (y_i - \theta^T x_i) \ .$$

In words, this cost is the *sum of squared errors.*

1. Letting

$$A \triangleq \begin{bmatrix} — x_1^T — \\ — x_2^T — \\ \vdots \\ — x_N^T — \end{bmatrix} \qquad Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}$$

we can rewrite $J(\theta)$ from above as

$$J(\theta) = \|Y - A\theta\|_2^2 \ .$$

show that

$$\theta^* = \arg\min_{\theta} J(\theta) = (A^T A)^{-1} A^T Y$$

Hint: Use the fact that $\|x\|_2^2 = x^T x$.

Solution:

Notice that

$$J = \theta^T A^T A\theta - 2Y^T A\theta + Y^T Y$$

Differentiate with respect to θ and set equal to zero:

$$\nabla_\theta J = 2A^T A - 2A^T Y = 0 \qquad \Rightarrow \theta^* = (A^T A)^{-1} A^T Y$$

Check that this is a minimum by obtaining the second derivative of J, and making sure it is positive

$$\nabla_\theta^2 J = A^T A > 0$$

2. Suppose the measurements arrive sequentially. Let $\hat{\theta}_N$ be the least-squares solution obtained front the first N measurements. We obtain a new measurement, and we want to use this information to compute $\hat{\theta}_{N+1}$ without recomputing the solution from scratch. We can do this via the 'Recursive Least Squares Filter'. Let

$$P_N = A_N^T A_N \qquad \text{where} \quad A_N = \begin{bmatrix} — x_1^T — \\ — x_2^T — \\ \vdots \\ — x_N^T — \end{bmatrix}$$

Show that the recursion

$$\hat{\theta}_N = \hat{\theta}_{N-1} + P_N^{-1} x_N (y_N - x_N^T \hat{\theta}_{N-1})$$

with

$$P_N^{-1} = P_{N-1}^{-1} - P_{N-1}^{-1} x_N (1 + x_N^T P_{N-1}^{-1} x_N)^{-1} x_N^T P_{N-1}^{-1}$$

produces the least squares solution using available measurements. That is, show that the formula from part 1 and this recursion agree for any N.

Solution:

The recursion for P_N is proved by applying the matrix inversion lemma with $A = P_{N-1}$, $B = x_N$, $C = 1$, $D = x_N^T$.

The recursion for $\hat{\theta}_N$ is obtained by using the fact that

$$P_m \hat{\theta}_m = A_m^T Y_m$$

for any m.

Expanding at step N, we get

$$\begin{aligned} P_N \hat{\theta}_N &= A_N^T Y_N \\ &= A_{N-1}^T Y_{N-1} + x_N y_N \\ &= P_{N-1} \hat{\theta}_{N-1} + x_N y_N \\ &= (P_N - x_N x_N^T) \hat{\theta}_{N-1} + x_N y_N \\ &= P_N \hat{\theta}_{N-1} + x_N (y_N - x_N^T \hat{\theta}_{N-1} \end{aligned}$$

so that

$$\hat{\theta}_N = \hat{\theta}_{N-1} + P_N^{-1} x_N (y_N - x_N^T \hat{\theta}_{N-1})$$

3. In 3050 A.D. NASA sends a spacecraft light-years away, towards a habitable planet. The spacecraft enters the planet's atmosphere, and the astronauts on board realize that they have no idea what the atmospheric drag coefficient (c_{atm}) or the acceleration due to gravity (g) of this new planet is. This would help them plan a safe and smooth descent trajectory to the surface. The spacecraft is known to fall according to

$$m_s\ddot{z} - c_{atm}|\dot{z}| = -m_s g$$

where $z(t)$ is the spacecraft's altitude.

(i) In the first scenario, the people on board never solved part 2 above, and so they crash-land (don't worry, everyone survives). They collected measurements of $\ddot{z}$ and $|\dot{z}|$ throughout the fall, and their data is available as `data.m` (the first column is $\ddot{z}$ and the second column is $|\dot{z}|$.) Obtain least-squares estimates for g and c_{atm}/m_s using the result from part 1.
Hint: Pick one of the measured values to be y, and the other to be part of x (you need to add something to the end of each x to account for a constant offset).

Solution:

```
%Pseudo-Code by KRISHNA SHANKAR
load data

X = [data(:,2), ones(length(data(:,1)),1)];
Y = data(:,1);

%One Step Least Squares Solution
v = (X'*X)^-1*X'*Y;

%Recursive Least Squares Solution
theta = zeros(2,1);
R = eye(2)*10^6;
for i = 1:length(X(:,1))
    x = [X(i,1);1];
    R = R-R*x*(1+x'*R*x)^(-1)*x'*R;
    Rst(i) = R(1,1);
    theta = theta+R*x*(Y(i)-x'*theta);
    Theta(:,i) = theta;
end

plot(Theta(2,:));
xlabel 'N'
ylabel 'Estimate of g'
```

We find that $c_{atm} \approx 5.15$ and $g \approx -20.24$.

4. In the second scenario, the people on board solved part 2, and they decide to use the RLS filter – Implement the recursion from part 2 (set $P_0 = \mathbb{I}_{n\times n} \times 10^6$ and $\hat{\theta}_0 = \vec{0}$). Plot your estimates of g and c_{atm}/m_s as a function of N, where N ranges from 1 to

number of measurements. Do your estimates converge to the answers from the previous question?

Solution: Below is the plot of the gravity estimate versus iteration

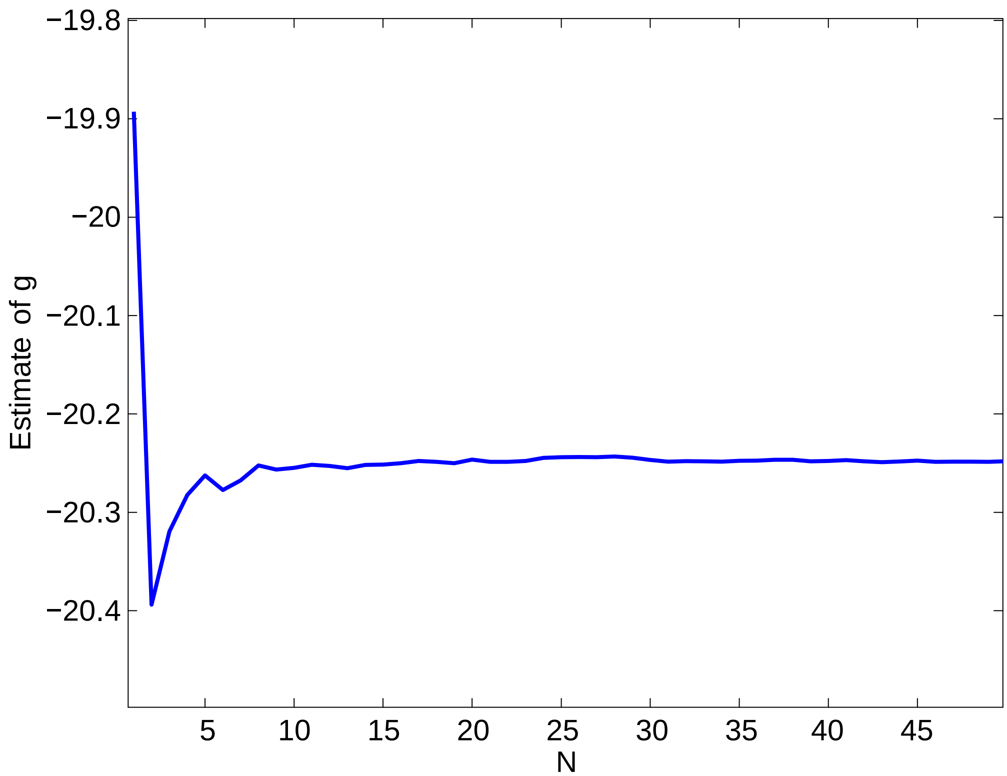

Figure 1: Estimates as a function of step N

Problem #1: Kidnapped Robot Problem

In this problem, you were asked to consider a simplified 1-dimensional model of a robot which has to estimate it's location using low-rate GPS measurements and higher rate accelerometer.

Part (a): You were asked to discretize the dynamic model:

$$m\frac{d^2x}{dt^2} + b\frac{dx}{dt} = u$$

where m is the vehicle mass (100 kg), and b is the drag constant, specfied as $b/m = -0.01s^{-1}$. Let's first divide the equation through by the mass, m:

$$\ddot{x} + \frac{b}{m}\dot{x} = \frac{u}{m} \tag{1}$$

There are many different types of discretized models that will provide good performance in this application. In general, we must find a first-order model that includes acceleration, $\ddot{x}$, as one of the states–since we must have a measurement equation that includes acceleration as a measurement.

A common approach is to somewhat ignore the vehicle dynamics and use a simple model based on integration of vehicle acceleration over a "short" period δT:

$$\begin{bmatrix} x_{k+1} \\ \dot{x}_{k+1} \\ \ddot{x}_{k+1} \end{bmatrix} = \begin{bmatrix} 1 & \delta T & \frac{(\delta T)^2}{2} \\ 0 & 1 & \delta T \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_k \\ \dot{x}_k \\ \ddot{x}_k \end{bmatrix} \tag{2}$$

where x_k is the vehicle position at t_k, $\dot{x}_k$ is the vehicle velocity at t_k, and $\ddot{x}_k$ is the vehicle acceleration at time t_k. This model makes the short-cut of assuming that vehicle acceleration is constant.

A slightly more accurate first-order continuous model can be derived by using Equation (1) and its derivative:

$$\frac{d^3x}{dt^3} = -\frac{b}{m}\ddot{x} + \frac{\dot{u}}{m} \tag{3}$$

which leads to the following first-order continuous model

$$\frac{d}{dt}\begin{bmatrix} x \\ \dot{x} \\ \ddot{x} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & -b/m & 0 \\ 0 & 0 & -b/m \end{bmatrix} \begin{bmatrix} x \\ \dot{x} \\ \ddot{x} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ (1/m) & 0 \\ 0 & (1/m) \end{bmatrix} \begin{bmatrix} u \\ \dot{u} \end{bmatrix} \tag{4}$$

If Equation (4) has the form $\dot{z} = A^c z + B^c u^c$, with $z = \begin{bmatrix} x & \dot{x} & \ddot{x} \end{bmatrix}^T$, then recall that for a constant input over the interval $[t_k, t_{k+1}]$, the discretized equation can be found from the convolution integral:

$$z_{k+1} = e^{A^c \delta T} z_k + \int_{t_k}^{t_{k+1}} e^{A^c \delta T} B^c u^c dt \triangleq A z_k + B u_k \ .$$

As we shall see below, the matrix B is irrelevant for this problem, and so we need only compute the matrix A:

$$A = e^{A^c \delta T} \ .$$

Since the accelerometer samples at 20 Hz ($\delta T = 0.05$ seconds), and the GPS sampling interval ($\delta T = 0.5$ seconds) is an integer multiple, we can choose $\delta T = 0.05$ for the discretization. Using the given problem data of $(b/m) = 0.01$, we have that (using the `expm` function in MATLAB):

$$A = \exp\left\{ 0.5 * \begin{bmatrix} 0 & 1 & 0 \\ 0 & -0.01 & 0 \\ 0 & 0 & -0.01 \end{bmatrix} \right\} = \begin{bmatrix} 1 & 0.05 & 0 \\ 0 & 0.9995 & 0 \\ 0 & 0 & 0.9995 \end{bmatrix}$$

Note that this exact discretization is surprisingly close to the crude model of Equation (2).

Part(b): Since there are no disturbances, except for the initial uncertainty of the vehicle's position, the Kalman filter dynamic update is quite simple:

$$\begin{aligned} \hat{x}_{k+1|k} &= A\hat{x}_k + Bu_k && (5) \\ P_{k+1|k} &= AP_{k|k}A^T. && (6) \end{aligned}$$

The covariance part of the measurement update takes the form:

$$P_{k+1|k+1} = P_{k+1|k} - P_{k+1|k}H^T(HP_{k+1|k}H^T + R)^{-1}HP_{k+1|k} \ . \quad (7)$$

where H is the measurement matrix and R is the variance of the measurement noise. During most measurement cycles, only the acceleromater measurement is available, and in that case, the measurement matrix and noise are:

$$\begin{aligned} H_{accel} &= \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \\ R_{accel} &= (0.0981 msec^{-2})^2 \end{aligned}$$

When both the GPS and accelerometers signal are available, then

$$\begin{aligned} H_{accel,GPS} &= \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \\ R_{accel} &= \begin{bmatrix} (0.0981 msec^{-2})^2 & 0 \\ 0 & (5.0m)^2 \end{bmatrix} \end{aligned}$$

The calculation is initialized with P_0, the initial uncertainty of the system state. If we assume that the initial uncertainties in position, velocity, and acceleration are mutually uncorrelated, then P_0 takes the form:

$$P_0 = \begin{bmatrix} P^0_{xx} & 0 & 0 \\ 0 & P^0_{vv} & 0 \\ 0 & 0 & P^0_{aa} \end{bmatrix}$$

where P^0_{xx} is the initial variance in position (which is assumed to 100 m^2), P^0_{vv} is the initial variance in vehicle velocity (which was not specified), and P^0_{aa} is the initial uncertainty in acceleration (which we can assume is $(0.0981\ m\ sec^{-1})^2$, the variance in the accelerometer noise). If you assume that the vehicle is at rest in its initial configuration, with no uncertainty, then $P^0_{vv} = 0$. A plot of position variance vs. k is shown in Figure 2(a). If the initial uncertainty in velocity is non-zero ($P_{vv} = 1.0$ in this example), then one obtains the position variance vs. k curve seen in Figure 2(b).

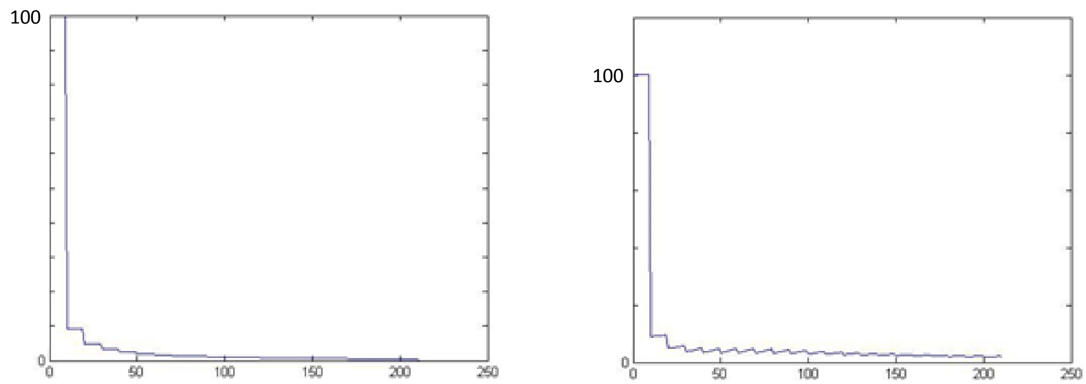

Figure 2: Plot of vehicle Position variance vs. k. (a) case where $P_{vv}^0 = 0$; (b) case where $P_{vv}^0 = 1$;