Solution 2: Least Squares

Problem 1:

Consider the least-squares (LS) estimation problem:

$$Y = \Phi\theta + \epsilon,$$

where Φ is the regressor matrix and θ is the parameter vector to be estimated

$$Y := \begin{bmatrix} y(0) \\ \vdots \\ y(N-1)) \end{bmatrix}, \qquad \Phi := \begin{bmatrix} \varphi^\top(0) \\ \vdots \\ \varphi^\top(N-1) \end{bmatrix}, \qquad \theta := \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_p \end{bmatrix}.$$

Assume that the noise, ϵ, is zero-mean Gaussian and correlated with $E\left\{\epsilon\epsilon^\top\right\} = R$. In this exercise we look for a linear estimator $\hat{\theta}$ of the form,

$$\hat{\theta} = Z^\top Y, \tag{2.1}$$

which is unbiased and minimizes its variance. For a given Φ show the following:

1. For a linear estimator of the form (2.1) to be unbiased we require that $Z^\top\Phi = I$.
2. The covariance matrix of any linear unbiased estimator of the form (2.1) is $\operatorname{cov}\left\{\hat{\theta}\right\} = Z^\top R Z$.
3. The covariance matrix of the best linear unbiased estimator (BLUE) $\hat{\theta}_Z$ with $\hat{\theta}_Z = \left(\Phi^\top R^{-1}\Phi\right)^{-1}\Phi^\top R^{-1} Y$ is $\operatorname{cov}\left\{\hat{\theta}_Z\right\} = \left(\Phi^\top R^{-1}\Phi\right)^{-1}$.
4. The best linear unbiased estimator $\hat{\theta}_Z$ exhibits the smallest variance in the class of all unbiased estimators, i.e. $\operatorname{cov}\left\{\hat{\theta}_Z\right\} \leq \operatorname{cov}\left\{\hat{\theta}\right\}$.

Hint: All covariance matrices are positive semi-definite and in our case we can assume that R is positive definite. The inverse of a positive definite matrix is also positive definite.

Solution

1. For a linear estimator of the form (2.1) to be unbiased we require that

$$\theta = \mathbb{E}\left\{\hat{\theta}\right\}.$$

Hence, for zero-mean Gaussian noise, ϵ, and fixed Φ we get

$$\theta = \mathbb{E}\left\{Z^\top Y\right\} = \mathbb{E}\left\{Z^\top\left(\Phi\theta + \epsilon\right)\right\} = Z^\top \Phi\theta,$$

which requires that

$$Z^\top \Phi = I.$$

2. The covariance matrix of any linear unbiased estimator of the form (2.1) is

$$\begin{aligned}
\operatorname{cov}\left\{\hat{\theta}\right\} &= \mathbb{E}\left\{\left(\hat{\theta}-\theta\right)\left(\hat{\theta}-\theta\right)^\top\right\} && \left[\hat{\theta} = Z^\top Y\right]\\
&= \mathbb{E}\left\{\left(Z^\top Y-\theta\right)\left(Z^\top Y-\theta\right)^\top\right\} && \left[Y = \left(\Phi\theta+\epsilon\right)\right]\\
&= \mathbb{E}\left\{\left(Z^\top\left(\Phi\theta+\epsilon\right)-\theta\right)\left(Z^\top\left(\Phi\theta+\epsilon\right)-\theta\right)^\top\right\} && \left[Z^\top\Phi = I\right]\\
&= Z^\top \mathbb{E}\left\{\epsilon\epsilon^\top\right\} Z && \left[\mathbb{E}\left\{\epsilon\epsilon^\top\right\} = R\right]\\
&= Z^\top R Z.
\end{aligned}$$

3. The covariance matrix of the best linear unbiased estimator (BLUE) $\hat{\theta}_Z$ with $\hat{\theta}_Z = \left(\Phi^\top R^{-1}\Phi\right)^{-1}\Phi^\top R^{-1} Y$ is

$$\begin{aligned}
\operatorname{cov}\left\{\hat{\theta}_Z\right\} &= Z^\top R Z && \left[Z^\top = \left(\Phi^\top R^{-1}\Phi\right)^{-1}\Phi^\top R^{-1}\right]\\
&= \left(\Phi^\top R^{-1}\Phi\right)^{-1}\Phi^\top R^{-1} R R^{-1}\Phi\left(\Phi^\top R^{-1}\Phi\right)^{-1}\\
&= \left(\Phi^\top R^{-1}\Phi\right)^{-1} && .
\end{aligned}$$

4. For $\hat{\theta}_Z$ to be the BLUE in the class of all unbiased estimators, we require that $\operatorname{cov}\left\{\hat{\theta}\right\} - \operatorname{cov}\left\{\hat{\theta}_Z\right\} \geq 0$:

$$\begin{aligned}
\operatorname{cov}\left\{\hat{\theta}\right\} - \operatorname{cov}\left\{\hat{\theta}_Z\right\} &= Z^\top R Z - \left(\Phi^\top R^{-1}\Phi\right)^{-1} && \left[Z^\top\Phi = I\right]\\
&= Z^\top R Z - Z^\top\Phi\left(\Phi^\top R^{-1}\Phi\right)^{-1}\Phi^\top Z\\
&= Z^\top\left[R - \Phi\left(\Phi^\top R^{-1}\Phi\right)^{-1}\Phi^\top\right] Z && \left[F = R - \Phi\left(\Phi^\top R^{-1}\Phi\right)^{-1}\Phi^\top\right]\\
&= Z^\top F Z.
\end{aligned}$$

Hence, we need to show that F is positive semidefinite so that $Z^\top F Z$ is positive semidefinite as well. We can show that

$$\begin{aligned}
F^\top R^{-1} F &= \left(R - \Phi\left(\Phi^\top R^{-1}\Phi\right)^{-1}\Phi^\top\right)^\top R^{-1}\left(R - \Phi\left(\Phi^\top R^{-1}\Phi\right)^{-1}\Phi^\top\right)\\
&= \left(R - \Phi\left(\Phi^\top R^{-1}\Phi\right)^{-1}\Phi^\top\right)\left(I - R^{-1}\Phi\left(\Phi^\top R^{-1}\Phi\right)^{-1}\Phi^\top\right)\\
&= R - \Phi\left(\Phi^\top R^{-1}\Phi\right)^{-1}\Phi^\top - \Phi\left(\Phi^\top R^{-1}\Phi\right)^{-1}\Phi^\top +\\
&\quad \Phi\left(\Phi^\top R^{-1}\Phi\right)^{-1}\Phi^\top R^{-1}\Phi\left(\Phi^\top R^{-1}\Phi\right)^{-1}\Phi^\top\\
&= R - \Phi\left(\Phi^\top R^{-1}\Phi\right)^{-1}\Phi^\top\\
&= F.
\end{aligned}$$

Since R is positive definite, R^{-1} is positive definite, we can conclude that $F = F^\top R^{-1} F$ has to be positive semidefinite and therefore cov $\left\{\hat{\theta}_Z\right\} \leq$ cov $\left\{\hat{\theta}\right\}$. Hence, $\hat{\theta}_Z$ is the BLUE with the smallest variance, but the assembly of Z requires knowledge of the error variance, R.

Problem 2:

Show that the following problems can be solved using linear least squares, with the estimate given by

$$\hat{\theta} = (X^T X)^{-1} X^T Y. \tag{2.2}$$

For each write down $\hat{\theta}$, X, and Y.

1. fitting the coeficients of a polynomial of degree k, $p(x) = \sum_{i=0}^{k} a_i x^i$, to n i.i.d. observations.
2. fitting the equation of a circle, $(x - x_c)^2 + (y - y_c)^2 = r^2$, to n i.i.d. observations.

Hint 1: Write the expression in matrix form, with the coefiecients of the polynomial isolated in a vector.

Hint 2: The general equation for a conic section is $F(x, y) = ax^2 + bxy + cy^2 + dx + ey + f = 0$, with a, b and c not all zero. A circle is a conic section.

Solution

The least squares solution is:

$$\min_{\hat{a}} ||Y - \hat{Y}||_2^2 = \min_{\hat{a}} J(\hat{a}) \tag{2.3}$$

$$\begin{aligned}
J(\hat{a}) &= (Y - \hat{Y})^T (Y - \hat{Y}) = \\
&= Y^T Y - 2Y^T \hat{Y} + \hat{Y}^T \hat{Y} = \\
&= Y^T Y - 2Y^T X\hat{a} + \hat{a}^T X^T X\hat{a} \\
\nabla_{\hat{a}} J &= -2Y^T X + 2\hat{a}^T X^T X = 0 \\
\hat{a} &= (X^T X)^{-1} X^T Y
\end{aligned}$$

1.

$$\hat{\theta} := \begin{bmatrix} \hat{a}_0 \\ \vdots \\ \hat{a}_k \end{bmatrix}, \qquad Y := \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \qquad X := \begin{bmatrix} 1 & x_1 & \dots & x_1^k \\ & \vdots & & \\ 1 & x_n & \dots & x_n^k \end{bmatrix}$$

2. Taking the equation of a circle and tranforming it into a general conic session:

$$1\ x^2 + 0\ xy + 1\ y^2 + (-2x_c)\ x + (-2\ y_c)\ y + (x_c^2 + y_c^2 - r^2) = 0. \tag{2.4}$$

Let $d = -2x_c$, $e = -2y_c$, $f = x_c^2 + y_c^2 - r^2$:

$$dx + ey + f = -x^2 - y^2. \tag{2.5}$$

$$\hat{\theta} := \begin{bmatrix} \hat{d} \\ \hat{e} \\ \hat{f} \end{bmatrix}, \qquad Y := \begin{bmatrix} -x_1^2 - y_1^2 \\ \vdots \\ -x_n^2 - y_n^2 \end{bmatrix}, \qquad X := \begin{bmatrix} x_1 & y_1 & 1 \\ & \vdots & \\ x_n & y_n & 1 \end{bmatrix}$$

Once $\hat{\theta}$ is found: $\hat{x_c} = -\frac{1}{2}\hat{d}$, $\hat{y_c} = -\frac{1}{2}\hat{e}$, $\hat{r} = \pm\sqrt{\hat{x_c}^2 + \hat{y_c}^2 - \hat{f}}$

MATLAB **Exercise:**

Problem 3:

An astronaut is in the vicinity of Saturn when a technical problem forces an automated emergency landing in a nearby moon. Unfortunately the space positioning system (SPS) of the spaceship is damaged. The color of the surface indicates this is either Iapetus ($g_I = 0.223 \pm 0.001$ ms^{-2}) or Rhea ($g_R = 0.264 \pm 0.001$ ms^{-2}). The values presented for the acceleration are in the format $g = \mu \pm \sigma$, i.e. the value of g follows a normal distribution with mean μ and standard deviation σ: $g \sim \mathcal{N}(\mu, \sigma^2)$.

Using a metric scale printed on the spaceship, a watch and a SysID book, an experiment is conducted to determine the local gravity at the surface, the results were recorded in the file "experiment1.dat". This file contains three columns: 1$^{\text{st}}$) the time since the beginning of the experiment in seconds, 2$^{\text{nd}}$) the measured position of the object in meters, 3$^{\text{rd}}$) the standard deviation of the measurement noise in meters.

Recall that the position of a SysID book moving in a gravitational field is given by

$$y(t) = y_0 + v_0 t + \frac{1}{2} a t^2 \tag{2.6}$$

1. Determine the local gravity on the surface using Least Squares. Plot the experimental data points along with the fit. Hint: Estimate the vector $\theta = [y_0, v_0, \frac{1}{2}a]$.
2. Provide the standard deviation of the estimates for $[y_0, v_0, a]$. What moon is this one?
3. After closer inspection the astronaut found that the scale had been damaged by the landing, so the measurements were more accurate in some parts of the scale than others. The results from the experiment were updated with the information of the non constant uncertainty of each measurement (see file "experiment2.dat"). Determine the local gravity on the surface using weighted Least Squares. Plot the experimental data points along with the fit.
4. Provide the standard deviation of the new estimates for $[y_0, v_0, a]$. What moon is this one?

Solution

1 and 2

In order to fit a polynomial using linear least squares, one can use the results from Problem 2. In matlab you can use the operator "\".

In order to estimate the error, one can use the result from Problem 1.2:

$$cov(\hat{\theta}) = Z^T RZ, \tag{2.7}$$

where R is a diagonal matrix, computed from the given standard deviations given in the file "experiment1.dat".

Using the a piori error estimates: $y_0 = 0.393 \pm 0.079$ m, $v_0 = 1.249 \pm 0.041$ m s^{-1}, $a_0 = -0.2736 \pm 0.0087$ m s^{-2}.

(Using errors estimated with reduced chi-square χ^2_μ, one obtains the "standard errors" of the mean estimate: $y_0 = 0.393 \pm 0.046$ m, $v_0 = 1.249 \pm 0.024$ m s^{-1}, $a_0 = -0.2736 \pm 0.0051$ m s^{-2}.)

The plot is show in Fig. 2.1.

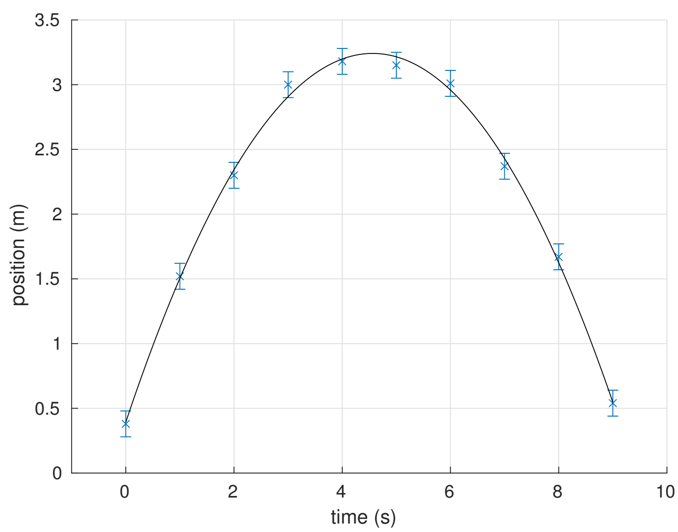

Figure 2.1: Least Squares Fit. $\chi^2_\nu = 0.335$

3 and 4

Using the information on "experiment2.dat", one can contruct the covariance matrix of the measurement noise. This is also the matrix used for the weighting.

Using the a piori error estimates: $y_0 = 1.133 \pm 0.077$ m, $v_0 = 0.923 \pm 0.036$ m s^{-1}, $a_0 = -0.2050 \pm 0.0078$ m s^{-2}

(Using error estimated with reduced chi-square χ^2_μ, one obtains the "standard errors" of the mean estimate: $y_0 = 1.13 \pm 0.22$ m, $v_0 = 0.92 \pm 0.10$ m s^{-1}, $a_0 = -0.205 \pm 0.023$ m s^{-2}.)

The plot is show in Fig. 2.2.

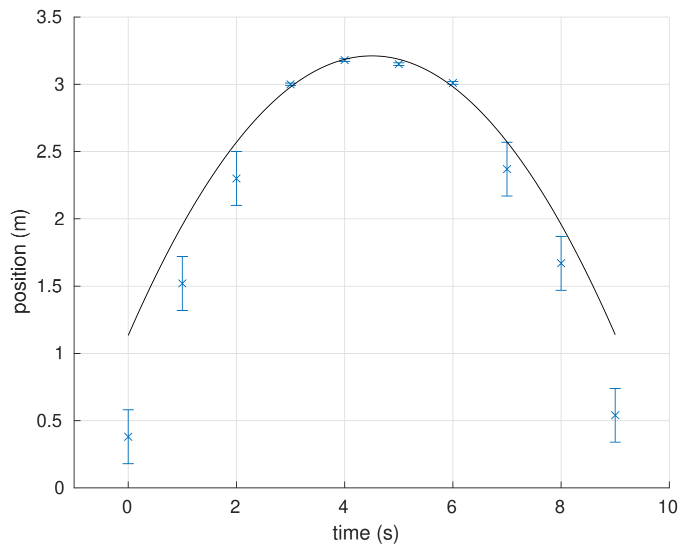

Figure 2.2: Weighted Least Squares Fit. $\chi^2_\nu = 8.304$

Problem 4:

A disk falls above the surface of a planet with constant rotational speed. The disk contains a sensor placed at a distance r from its centre, measuring a quantity correlated with the height of the disk. Figure 2.3 shows the apparatus.

The output of the sensor is given by:

$$y(t) = z_0 + v_0 t + \frac{1}{2} g t^2 + r \cos(\theta_0 + \omega_0 t) + \epsilon, \tag{2.8}$$

where $\omega_0 = 10$ rad s^{-1} is the angular velocity, $\theta_0 = \frac{\pi}{8}$ rad is the initial angle, r is the unknown distance of the sensor to the centre of the disk, z_0 is the unknown initial height of the centre of the disk, v_0 is the unknow initial velocity of the centre of the disk, t is the time since the beggining of the experiment, g is the unknown gravity constant at the surface of the planet, and $\epsilon \sim \mathcal{N}(0, \sigma)$ the normally distributed noise.

Given the data provided in "experiment3.dat", and using linear least squares, estimate the values for the variables $\{r, z_0, v_0, g\}$.

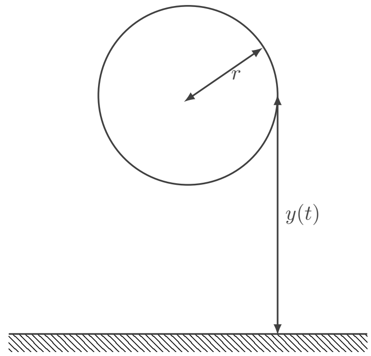

Figure 2.3: Experimental apparatus scheme

Hint: Write an expression where the output depends linearly on the least squares parameters.

Solution

E.g. the following

```
%Given parameters
w0 = 10
a0 = pi/8

%Import data
experiment1 = importdata("experiment3.dat").data;
t = experiment1(:,1);
y = experiment1(:,2);

plot(t, y)

%Build auxiliary matrix for linear least squares
X = [1+0*t, t, (1/2)*t.^2, cos(a0 + w0*t)];

%Least squares solution
sol = X \ y;

disp("Least Squares solution")
disp(sol')

true_val = [12.42, 44.19, -6.42, 1.53];
disp("True values")
disp(true_val)
```