

Big Data Analytics - Spring 2014 - Homework #3

1 (This problem is a particular case of Problem 3. So, if you feel confident you can skip it and then extract answers from Problem 3).

Consider the following set of measurements of the unknown variables x_1 and x_2 :

$$y_1 = x_1 + x_2 + \nu_1,$$

$$y_2 = x_1 - x_2 + \nu_2,$$

$$y_3 = -x_1 + x_2 + \nu_3,$$

where y_i are measurement results, and ν_i represent random error of measurement and are independent identically distributed (i.i.d.) with zero mean and variance σ^2 :

$$\mathbf{E}\nu_i, \quad \mathbf{E}\varepsilon_i^2 = \sigma^2, \quad i = 1, 2, 3.$$

(a) Write it in matrix form

$$y = Ax + \nu$$

and write the matrices A and $S = \text{Var}(\nu)$.

(b) Find the variance matrix $\text{Var}(\hat{x})$ for the optimal linear estimate of x and variances of \hat{x}_1 and \hat{x}_2 .

2 Consider two measurements of one unknown variable x with the correlated noise. Specifically, suppose that

$$y_1 = x + \nu_1,$$

$$y_2 = x + \nu_2,$$

where

$$\nu_1 = \varepsilon_1 + \varepsilon_0,$$

$$\nu_2 = \varepsilon_2 + \varepsilon_0,$$

$$\varepsilon_1, \varepsilon_2 \sim (0, \sigma_1^2), \quad \varepsilon_0 \sim (0, \sigma_0^2),$$

$$\sigma_0^2 + \sigma_1^2 = \sigma^2, \quad r = \frac{\sigma_0^2}{\sigma_0^2 + \sigma_1^2}.$$

- (a) Same as in 1.
- (b) Same as in 1.
- (c) Analyze how the variance of \hat{x} depends on the correlation parameter r for $0 \leq r \leq 1$. Is higher correlation good or bad for estimation in this example? A graph might be helpful. How would you explain such behavior?

3 Consider the same measurement scheme as in Problem 1, but with the correlated noise. Specifically, suppose that

$$\nu_1 = \varepsilon_1 + \varepsilon_0,$$

$$\nu_2 = \varepsilon_2 + \varepsilon_0,$$

$$\nu_3 = \varepsilon_3 + \varepsilon_0,$$

$$\varepsilon_1, \varepsilon_2, \varepsilon_3 \sim (0, \sigma_1^2), \quad \varepsilon_0 \sim (0, \sigma_0^2),$$

$$\sigma_0^2 + \sigma_1^2 = \sigma^2, \quad r = \frac{\sigma_0^2}{\sigma_0^2 + \sigma_1^2}.$$

- (a) Same as in 1.
- (b) Same as in 1.
- (c) Analyze how the variances of \hat{x}_1 and \hat{x}_2 depend on the correlation parameter r for $0 \leq r \leq 1$. Is higher correlation good or bad for estimation in this example? A graph might be helpful. How would you explain such behavior?

Feel free to use symbolic packages (e.g., Symbolic Toolbox in MatLab).