Big Data Analytics - Spring 2014 - Project 1
Calibration Problem

In this project you will implement and demonstrate optimal calibration for a linear estimation. The underlying process imitates a simple signal measurement experiment and is heavily based on Homework 4. Here are suggested phases of the project.

1 Measurement Simulation:

(a) Randomly generate some "unknown" profile $x \sim (0, F)$.

(b) Create a matrix A.

(c) Simulate a measurement

$$y = Ax + \nu, \quad \nu \sim (0, \sigma^2 I).$$

2 Calibration (assuming that A is unknown):

(a) Randomly generate calibration signals $\varphi_k, \ k = 1, \ldots, K$ in the same way as you generated x.

(b) Simulate calibration measurements

$$\psi_k = A\varphi_k + \nu_k.$$

(c) Collect canonical calibration information (G, H).

(d) Compute A_0 (an estimate of A) and J.

3 Using your simulated observation y construct an optimal linear estimate $\widehat{x}$ and the variance matrix $\mathrm{Var}(\widehat{x} - x)$. Show on the same graph:

(a) The original signal x (a curve with components x_i),

(b) Its estimate $\widehat{x}$ (a curve with components $\widehat{x}_i$),

(c) Standard deviations for the estimates $\widehat{x}_i$

$$\sqrt{\mathsf{E}||\widehat{x}_i - x_i||^2} = \sqrt{\mathrm{Var}(\widehat{x} - x)_{ii}} = \sqrt{Q_{ii}}$$

can be illustrated by showing the corresponding "corridor" around $\widehat{x}_i$).

4 Illustrate estimation (Phase 3) when you not only increase the number of calibration measurements K but also measure the original signal x N times:

(a) Simulate N measurements $y_n = Ax + \nu_n$ for $n = 1, \ldots, N$ and collect the appropriate information.

(b) Simulate K calibration measurements, collect canonical calibration information, and compute A_0 and J.

(c) Using these two types of information construct and show (as in Phase 3) an optimal linear estimate $\widehat{x}$ and its precision.

(d) Using these two types of information construct and show (as in Phase 3) an optimal linear estimate $\widehat{x}$ and its precision.

(e) Do the above for several cases with numbers N and K "small" "medium", and "large".

(f) Show how estimation precision depends on N and K. To do that you could show total estimation error

$$\mathsf{E}||\widehat{x} - x||^2 = \mathrm{tr}Q$$

as a function of N and K. To illustrate a function of two variables you can show it, e.g., as a surface or as a pseudocolor image. It might be interesting to indicate contour lines (curve along which the function has a constant value).

Please do not hesitate to ask questions. For more details please see Lecture 12 (pages 3-8) and Lecture 8 (pages 8-11) and the sample code from `http://www.math.umt.edu/golubtsov/BD_2014_S/HW4_Demo/Signal_2.m`