EE263 homework 5

1. *Curve-smoothing.* We are given a function $F : [0,1] \to \mathbf{R}$ (whose graph gives a curve in $\mathbf{R}^2$). Our goal is to find another function $G : [0,1] \to \mathbf{R}$, which is a *smoothed* version of F. We'll judge the smoothed version G of F in two ways:

   - *Mean-square deviation from F,* defined as
   $$D = \int_0^1 (F(t) - G(t))^2 \, dt.$$
   - *Mean-square curvature,* defined as
   $$C = \int_0^1 G''(t)^2 \, dt.$$

   We want *both* D and C to be small, so we have a problem with two objectives. In general there will be a trade-off between the two objectives. At one extreme, we can choose $G = F$, which makes $D = 0$; at the other extreme, we can choose G to be an affine function (*i.e.*, to have $G''(t) = 0$ for all $t \in [0,1]$), in which case $C = 0$. The problem is to identify the optimal trade-off curve between C and D, and explain how to find smoothed functions G on the optimal trade-off curve. To reduce the problem to a finite-dimensional one, we will represent the functions F and G (approximately) by vectors $f, \ g \in \mathbf{R}^n$, where
   $$f_i = F(i/n), \quad g_i = G(i/n).$$
   You can assume that n is chosen large enough to represent the functions well. Using this representation we will use the following objectives, which approximate the ones defined for the functions above:

   - *Mean-square deviation,* defined as
   $$d = \frac{1}{n} \sum_{i=1}^{n} (f_i - g_i)^2.$$
   - *Mean-square curvature,* defined as
   $$c = \frac{1}{n-2} \sum_{i=2}^{n-1} \left(\frac{g_{i+1} - 2g_i + g_{i-1}}{1/n^2} \right)^2.$$

   In our definition of c, note that
   $$\frac{g_{i+1} - 2g_i + g_{i-1}}{1/n^2}$$
   gives a simple approximation of $G''(i/n)$. You will only work with this approximate version of the problem, *i.e.*, the vectors f and g and the objectives c and d.

   (a) Explain how to find g that minimizes $d + \mu c$, where $\mu \geq 0$ is a parameter that gives the relative weighting of sum-square curvature compared to sum-square deviation. Does your method always work? If there are some assumptions you need to make (say, on rank of some matrix, independence of some vectors, etc.), state them clearly. Explain how to obtain the two extreme cases: $\mu = 0$, which corresponds to minimizing d without regard for c, and also the solution obtained as $\mu \to \infty$ (*i.e.*, as we put more and more weight on minimizing curvature).

   (b) Get the file `curve_smoothing.m` from the course web site. This file defines a specific vector f that you will use. Find and plot the optimal trade-off curve between d and c. Be sure to identify any critical points (such as, for example, any intersection of the curve with an axis). Plot the optimal g for the two extreme cases $\mu = 0$ and $\mu \to \infty$, and for three values of μ in between (chosen to show the trade-off nicely). On your plots of g, be sure to include also a plot of f, say with dotted line type, for reference. Submit your Matlab code.

2. *Simultaneous left inverse of two matrices.* Consider a system where

$$y = Gx, \quad \tilde{y} = \tilde{G}x$$

where $G \in \mathbf{R}^{m\times n}$, $\tilde{G} \in \mathbf{R}^{m\times n}$. Here x is some variable we wish to estimate or find, y gives the measurements with some set of (linear) sensors, and $\tilde{y}$ gives the measurements with some *alternate* set of (linear) sensors. We want to find a *reconstruction matrix* $H \in \mathbf{R}^{n\times m}$ such that $HG = H\tilde{G} = I$. Such a reconstruction matrix has the nice property that it recovers x perfectly from *either* set of measurements (y or $\tilde{y}$), *i.e.*, $x = Hy = H\tilde{y}$. Consider the specific case

$$G = \begin{bmatrix} 2 & 3 \\ 1 & 0 \\ 0 & 4 \\ 1 & 1 \\ -1 & 2 \end{bmatrix}, \quad \tilde{G} = \begin{bmatrix} -3 & -1 \\ -1 & 0 \\ 2 & -3 \\ -1 & -3 \\ 1 & 2 \end{bmatrix}.$$

Either find an explicit reconstruction matrix H, or explain why there is no such H.

3. *Modifying measurements to satisfy known conservation laws.* A vector $y \in \mathbf{R}^n$ contains n measurements of some physical quantities $x \in \mathbf{R}^n$. The measurements are good, but not perfect, so we have $y \approx x$. From physical principles it is known that the quantities x must satisfy some linear equations, *i.e.*,

$$a_i^T x = b_i, \qquad i = 1, \ldots, m,$$

where $m < n$. As a simple example, if x_1 is the current in a circuit flowing into a node, and x_2 and x_3 are the currents flowing out of the node, then we must have $x_1 = x_2 + x_3$. More generally, the linear equations might come from various conservation laws, or balance equations (mass, heat, energy, charge …). The vectors a_i and the constants b_i are known, and we assume that $a_1, \ldots, a_m$ are independent. Due to measurement errors, the measurement y won't satisfy the conservation laws (*i.e.*, linear equations above) exactly, although we would expect $a_i^T y \approx b_i$. An engineer proposes to adjust the measurements y by adding a correction term $c \in \mathbf{R}^n$, to get an adjusted estimate of x, given by

$$y_{\text{adj}} = y + c.$$

She proposes to find the smallest possible correction term (measured by $\|c\|$) such that the adjusted measurements y_{adj} satisfy the known conservation laws. Give an explicit formula for the correction term, in terms of y, a_i, b_i. If any matrix inverses appear in your formula, explain why the matrix to be inverted is nonsingular. Verify that the resulting adjusted measurement satisfies the conservation laws, *i.e.*, $a_i^T y_{\text{adj}} = b_i$.

4. *General rank update formula* We proved a special case of the rank-1 update formula in class. For this problem, prove the more general rank update formula:

$$(A + UV)^{-1} = A^{-1} - A^{-1}U(I + VA^{-1}U)^{-1}VA^{-1}$$

and explain why such a formula might be computationally beneficial. You may assume that all matrices being inverted are invertible.