Recitation 12

November 12, 2015

Problem 1. Is the system $Ax = b$ consistent for

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \text{ and } b = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} ?$$

Find all vectors $\widehat{x} \in \mathbb{R}^3$ such that $A\widehat{x}$ is the closest to b vector in $Col(A)$.

Explain why you are not getting a unique $\widehat{x}$.

Solution:

Write the augmented matrix of the system $Ax = b$, and row reduce it. You get

$$\begin{bmatrix} 1 & 1 & 0 & |1 \\ 1 & 0 & 1 & |2 \\ 1 & 1 & 0 & |3 \\ 1 & 0 & 1 & |4 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 0 & |1 \\ 1 & 0 & 1 & |2 \\ 0 & 0 & 0 & |2 \\ 0 & 0 & 0 & |0 \end{bmatrix}$$

Since there is a pivot in the last column of augmented matrix, the system is inconsistent. Put simply, the 3rd row is the equation $0 = 2$, so of course it is inconsistent.

We use the normal equation $A^T A\widehat{x} = A^T b$ to find the vectors $\widehat{x}$ in the domain such that $A\widehat{x}$ is as close to b as possible, i.e. we are looking for $\widehat{x}$ such that $A\widehat{x}$ is exactly the projection $\widehat{b}$ onto $Col(A)$. The equation $A^T A\widehat{x} = A^T b$ reads

$$\begin{bmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{bmatrix} \widehat{x} = \begin{bmatrix} 10 \\ 4 \\ 6 \end{bmatrix}$$

Solving the system gives

$$\begin{bmatrix} 4 & 2 & 2 & |10 \\ 2 & 2 & 0 & |\ 4 \\ 2 & 0 & 2 & |\ 6 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 & |3 \\ 1 & 1 & 0 & |2 \\ 2 & 1 & 1 & |5 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 & |\ 3 \\ 0 & 1 & -1 & |-1 \\ 0 & 0 & 0 & |\ 0 \end{bmatrix}$$

So x_3 is a free variable, $x_2 = -1 + x_3$ and $x_1 = 3 - x_3$, and so the solutions $\widehat{x}$ are all vectors of the form

$$\begin{bmatrix} 3 \\ -1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

We didn't get a unique solution $\widehat{x}$ because the columns of A were linearly **dependent**, and so A, and therefore $A^T A$, have non-zero null-space. Thus the system $A^T A\widehat{x} = A^T b$ can't have unique solution. If you don't remember the conditions when the system has infinitely many solutions, then it's too bad, you need to review this stuff. But really, if v is non-zero, and s.t. $Av = 0$, then $A^T Av$ is also 0, and so for any solution $\widehat{x}_0$ of the system $A^T A\widehat{x} = A^T b$, $\widehat{x}_0 + cv$ is again a solution, for any scalar c.

Problem 2. Describe all least square solutions of the system

$$\begin{cases} x + y = 2 \\ x + y = 4 \end{cases}$$

Solution:

We do the same thing as before. First write the system of equations in the matrix form

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

The system is obviously inconsistent, so we run the least-squares process. The equation $A^T A\widehat{x} = A^T b$ in this case reads

$$\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \widehat{x} = \begin{bmatrix} 6 \\ 6 \end{bmatrix}$$

Solving the system, you get that y is a free variable, and $x = 3 - y$. In other words, the set of least-squares solutions is given by the equation $x + y = 3$, which if you look at the system, makes a lot of sense.
Note that here you also don't have a unique solution $\widehat{x}$, by the same reason as in Problem 1.

Problem 3. Suppose for the matrix A you know the result of orthonormalization of its columns, obtained by using Gramm-Schmidt. Use this data to obtain least-squares solution of the system $Ax = b$. The numbers are as follows:

$$A = \begin{bmatrix} 1 & -1 \\ 1 & 4 \\ 1 & -1 \\ 1 & 4 \end{bmatrix}, \; Q = \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \\ 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix}, \; b = \begin{bmatrix} -1 \\ 6 \\ 5 \\ 7 \end{bmatrix}$$

Solution:
To find R, you can use the formula $R = Q^T A$. So R is

$$R = \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ -1/2 & 1/2 & -1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 4 \\ 1 & -1 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 0 & 5 \end{bmatrix}$$

To find the least-squares solution $\widehat{x}$, we can use the formula $\widehat{x} = R^{-1}Q^T b$. So

$$\widehat{x} = \begin{bmatrix} 2 & 3 \\ 0 & 5 \end{bmatrix}^{-1} \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ -1/2 & 1/2 & -1/2 & 1/2 \end{bmatrix} \begin{bmatrix} -1 \\ 6 \\ 5 \\ 7 \end{bmatrix} = \begin{bmatrix} 29/10 \\ 9/10 \end{bmatrix}$$

Problem 4. Suppose you are observing some machine (I don't know, say, a magic box), and after random intervals of time this machine shows you a number[1].
At the times

$$x_1 = 1, \; x_2 = 1.5, \; x_3 = 2, \; x_4 = 2.5, \; x_5 = 3$$

(I agree, these four time intervals don't seems **that** random) the machine produced the following numbers:

$$y_1 = 1, \; y_2 = 1.5, \; y_3 = 2.5, \; y_4 = 4, \; y_5 = 5.5$$

You would like to predict what would be next, i.e. you are trying to model how the machine works[2] .

Find the least-squares line $y = \beta_0 + \beta_1 x$ approximating the work of the machine. What is the length of the error term (i.e. the length of the residual vector)?

Try to approximate work of this machine by a parabola $y = \beta_0 + \beta_1 x + \beta_2 x^2$ (i.e. assume that the data (x_i, y_i) occurs along a parabola). Compute the length of the error term in this case. So which approximation is better?

Solution:
Let's first do approximation by the line $y = \beta_0 + \beta_1 x$. Ideally, we would like this approximation to be exact. I mean, we would like our line $y = \beta_0 + \beta_1 x$ to pass through all the data points (x_i, y_i), namely, through $(1, 1)$, $(1.5, 1.5)$, $(2, 2.5)$, $(2.5, 4)$ and $(3, 5.5)$. Passing through a point means that the point

[1]Since it is a **magic** box, let's say it produces a rainbow in the sky, shaped as the number it outputs. If you prefer seeing **dark** magic, let's say the machine produces a number made of fire and blood of innocents. I don't know how it would work, I am not a magician.

[2]To motivate that, suppose if you guess what's next, you will win a bonus. If you prefer dark magic, let's say if you guess what's next, you will save a bunch of innocent innocents.

satisfies the equation $y = \beta_0 + \beta_1 x$. So ideally all of the following equations will be satisfied (ideally line passes through all of the points):

$$\begin{cases} \beta_0 + \beta_1 \cdot 1 & = 1 \\ \beta_0 + \beta_1 \cdot 1.5 & = 1.5 \\ \beta_0 + \beta_1 \cdot 2 & = 2.5 \\ \beta_0 + \beta_1 \cdot 2.5 & = 4 \\ \beta_0 + \beta_1 \cdot 3 & = 5.5 \end{cases}$$

Writing this system in the matrix form, you get

$$\begin{bmatrix} 1 & 1 \\ 1 & 1.5 \\ 1 & 2 \\ 1 & 2.5 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1.5 \\ 2.5 \\ 4 \\ 5.5 \end{bmatrix}$$

This system is inconsistent, which means that there is no line passing through all the data points. Not such a big surprise. But we can find least-squares solution, which would be the "best possible" approximation. In the previous notation,

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1.5 \\ 1 & 2 \\ 1 & 2.5 \\ 1 & 3 \end{bmatrix}, \widehat{x} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, b = \begin{bmatrix} 1 \\ 1.5 \\ 2.5 \\ 4 \\ 5.5 \end{bmatrix}$$

So you again have to solve the equations $A^T A\widehat{x} = A^T b$, and $\widehat{x} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$ will give you coefficients of the line you are looking for. Notice that here you will have a **unique** solution $\widehat{x}$. I am not going to compute this. Sorry...

The same idea for approximation by a parabola $y = \beta_0 + \beta_1 x + \beta_2 x^2$. Ideally, you would like parabola to go through all the data points, i.e. all the points would satisfy the equation $y = \beta_0 + \beta_1 x + \beta_2 x^2$. This gives

$$\begin{cases} \beta_0 + \beta_1 \cdot 1 + \beta_2 \cdot 1^2 & = 1 \\ \beta_0 + \beta_1 \cdot 1.5 + \beta_2 \cdot 1.5^2 & = 1.5 \\ \beta_0 + \beta_1 \cdot 2 + \beta_2 \cdot 2^2 & = 2.5 \\ \beta_0 + \beta_1 \cdot 2.5 + \beta_2 \cdot 2.5^2 & = 4 \\ \beta_0 + \beta_1 \cdot 3 + \beta_2 \cdot 3^2 & = 5.5 \end{cases}$$

In the matrix form, you get

$$\begin{bmatrix} 1 & 1 & 1^2 \\ 1 & 1.5 & 1.5^2 \\ 1 & 2 & 2^2 \\ 1 & 2.5 & 2.5^2 \\ 1 & 3 & 3^2 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1.5 \\ 2.5 \\ 4 \\ 5.5 \end{bmatrix}$$

You do the same thing with least-squares again.

Problem 5. Define an inner product on $\mathbb{P}_2$ by

$$\langle p, q \rangle = p(-1)q(-1) + p(0)q(0) + p(1)q(1)$$

Compute $\langle -2 + t + 2t^2, 3 - 2t \rangle$.

Compute the orthogonal projection of the polynomial $1 + t$ to the subspace spanned by $p = -2 + t + 2t^2, q = 3 - 2t$.

Solution:

Computing:

$$\langle -2 + t + 2t^2, 3 - 2t \rangle = -1 \cdot 5 + (-2) \cdot 3 + 1 \cdot 1 = -9$$

To orthogonalize $\{p = -2 + t + 2t^2, q = 3 - 2t\}$, use Gramm-Schmidt. Take $v_1 = -2 + t + 2t^2$, and compute

$$v_2 = (3 - 2t) - \frac{\langle 3 - 2t, -2 + t + 2t^2 \rangle}{\langle -2 + t + 2t^2, -2 + t + 2t^2 \rangle}(-2 + t + 2t^2) = (3 - 2t) + \frac{9}{6}(-2 + t + 2t^2)$$

So $v_2 = 3t^2 - 1/2t$. Ok, so now we've got orthogonal basis v_1, v_2 of $Span(p, q)$. Now you use the projection formula

$$proj_W(1+t) = \frac{\langle 1+t, -2+t+2t^2\rangle}{\langle -2+t+2t^2, -2+t+2t^2\rangle}(-2+t+2t^2) + \frac{\langle 1+t, 3t^2-1/2t\rangle}{\langle 3t^2-1/2t, 3t^2-1/2t\rangle}(3t^2-1/2t)$$

Please, do the calculation.

Problem 6. Prove that for any $n \times n$ invertible matrix A, the formula $\langle u, v\rangle := (Au) \cdot (Av) = (Au)^T(Av)$ defines an inner product on $\mathbb{R}^n$.
Hint:
You just need to check that this operation satisfies the definition of the inner product.