Recitation 12

November 12, 2015

Review

Projection to a subspace: if $W \subset \mathbb{R}^n$ is a subspace, and $y \in \mathbb{R}^n$ is a vector, to find $proj_W(y)$ you need to

- Find orthogonal basis $\{u_1, \ldots, u_m\}$ of W.
- Then

$$proj_W(y) = \frac{y \cdot u_1}{u_1 \cdot u_1} u_1 + \dots + \frac{y \cdot u_m}{u_m \cdot u_m} u_m$$

If you already know an orthoNORMAL basis v_1, \ldots, v_m of W: put $U = [v_1 \ldots v_m]$, then $proj_W(y) = UU^T y$.

Gramm-Schmidt: Let x_1, \ldots, x_m be a basis in $W \subset \mathbb{R}^n$. You want to find a new orthogonal basis using the old maybe-not-orthogonal one. You can use formulas:

$$v_{1} = x_{1}$$

$$v_{2} = x_{2} - \frac{x_{2} \cdot v_{1}}{v_{1} \cdot v_{1}} v_{1}$$

$$v_{3} = x_{3} - \frac{x_{3} \cdot v_{1}}{v_{1} \cdot v_{1}} v_{1} - \frac{x_{3} \cdot v_{2}}{v_{2} \cdot v_{2}} v_{2}$$

$$\dots = \dots$$

$$v_{m} = x_{m} - \frac{x_{m} \cdot v_{1}}{v_{1} \cdot v_{1}} v_{1} - \frac{x_{m} \cdot v_{2}}{v_{2} \cdot v_{2}} v_{2} - \dots - \frac{x_{m} \cdot v_{m-1}}{v_{m-1} \cdot v_{m-1}} v_{m-1}$$

QR-factorization: If x_1, \ldots, x_m are columns of A, apply Gramm-Schmidt, get v_1, \ldots, v_m . These would be columns of Q. Then $R = Q^T A$. This will ensure that A = QR, i.e. that's how you get the QR-factorization.

Least-squares: If a system Ax = b is inconsistent, we can find a **least-squares** solution, which is a solution to the equation $A^T A \hat{x} = A^T b$ (which is called the **normal equation**).

The **idea** is that \hat{x} is a vector in the domain of A such that $A\hat{x}$ is as close to b as possible. And so \hat{x} is "the best" substitute for a solution of Ax = b. This $A\hat{x}$ is just the projection of b onto Col(A).

If A has lin. independent columns and if you know the QR-decomposition of A, A = QR, then $\hat{\mathbf{x}} = \mathbf{R}^{-1}\mathbf{Q}^{T}\mathbf{b}$.

Inner product: inner product walks like the dot product, talks like the dot product, and behaves like the dot product. The only difference is, it's defined in a vector space which does not look like \mathbb{R}^n .

So to do things with inner product you use the same methods as for dot product, but replacing dot product by this inner product. For example, here is how the projection formula looks like for a projection of $y \in V$ in some inner product space V onto some subspace $W \subset V$ with fixed orthogonal basis $\{u_1, \ldots, u_m\}$:

$$proj_W(y) = \frac{\langle y, u_1 \rangle}{\langle u_1, u_1 \rangle} u_1 + \dots + \frac{\langle y, u_m \rangle}{\langle u_m, u_m \rangle} u_m$$

Problems

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 $\hat{x} \in \mathbb{R}^3$ such that $A\hat{x}$ is the closest to b vector in Col(A). Explain why you are not getting a unique \hat{x} .

Problem 2. Describe all least square solutions of the system

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

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:

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

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 number¹. 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 works². 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?

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.

(Hint: p, q are not orthogonal, so first you need to orthogonalize them)

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 .

¹Since 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.