Recitation 13

November 19, 2015

Problem 1. For the space $C[0,1]$ of continuous functions on the interval $[0,1]$ there is an inner product defined by

$$\langle f, g\rangle = \int_0^1 f(t)g(t)dt$$

Let $f(t) = t - 2$ and $g(t) = e^t$. Compute $\langle f, g\rangle$.

Solution. Ok, this is pretty straightforward. You just compute the integral (I hope I didn't mess up integration by parts)

$$\int_0^1 (t-2)e^t dt = -2e + 2 + e - \int_0^1 e^t dt = 2 - e - e + 1 = 3 - 2e$$

Problem 2. Suppose you and two friends of yours, call them A and B, are making a team for a class project. The project involves measurements of something, and then predicting how this something works. If you want, you can assume that you are trying to model how a magic box showing you random numbers works. You have agreed to do the modeling, and your friends are going to do the measurements.

Unfortunately, professor assigned another person to be on your team, namely, that weird guy named Q whom you know to be a bad guy. Maybe he punched a cat once, or maybe even something worse. But suppose you feel bad to just through away Q's measurements. So what you decide to do is, you are going to weight them half as much as the measurements that your friends did.

You are trying to do the approximation by a line $y = \beta_0 + \beta_1 x$. You friend A tells you that his data is $(x_1, y_1) = (1, 0.5)$. Your friend B gives you his data: it is $(x_2, y_2) = (1.5, 1.5)$. Now Q gives you his measurements. They are $(x_3, y_3) = (2, 3)$ and $(x_4, y_4) = (3, 5)$. You are suspicious: why did Q do two measurements instead of just one? What is he trying to prove? Weird...

Find the weighted least-squares line $y = \beta_0 + \beta_1 x$ approximating the data.

Solution. The weight matrix W will be the diagonal matrix with weights on the diagonal, i.e. $W = diag(1, 1, 1/2, 1/2)$. The secret is that we can actually re-scale W to get rid of the fractions. So really you can take W to be

$$W = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

If we did the usual least-squares solution, we would do the system of equations

$$\begin{cases} \beta_0 + \beta_1 \cdot 1 & = 0.5 \\ \beta_0 + \beta_1 \cdot 1.5 & = 1.5 \\ \beta_0 + \beta_1 \cdot 2 & = 3 \\ \beta_0 + \beta_1 \cdot 3 & = 5 \end{cases}$$

and so we are really trying to solve the system $Ax = y$ with

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1.5 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}, \; x = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, \; y = \begin{bmatrix} 0.5 \\ 1.5 \\ 3 \\ 5 \end{bmatrix}$$

However, since we agreed to weight different data differently, we need to re-scale the equation $Ax = y$ by W to obtain $WAx = Wy$. That's the thing we are going to try and solve. Note that if $Ax = y$ **had an honest**

solution, i.e. not just the least-squares one, then multiplying by W won't change the solution (and that's good!). The equation $WAx = Wy$ is

$$\begin{bmatrix} 2 & 2 \\ 2 & 3 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 3 \\ 5 \end{bmatrix}$$

Ok, now you just do the usual least-squares business. Multiply both sides of $WAx = Wy$ by $(WA)^T$, find its solution $\widehat{x} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$, that would be you **weighted** least-squares solution. You have

$$\begin{bmatrix} 2 & 2 & 1 & 1 \\ 2 & 3 & 2 & 3 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 2 & 3 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 1 & 1 \\ 2 & 3 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 3 \\ 5 \end{bmatrix}$$

which is

$$\begin{bmatrix} 10 & 15 \\ 15 & 26 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 16 \\ 32 \end{bmatrix}$$

Please, finish the calculation, it's too much for me ☺

Problem 3. Now we are going to do some of that trend analysis that they have now. You've seen it. Suppose we are doing measurements at the points $t = -5, -3, -1, 1, 3, 5$.
Show that the first three orthogonal polynomials are $p_0(t) = 1, p_1(t) = t, p_2(t) = \frac{3}{8}t^2 - \frac{35}{8}$.
Fit a quadratic trend function to the data

$$(-5, 1),\ (-3, 1),\ (-1, 4),\ (1, 4),\ (3, 6),\ (5, 8)$$

Solution. You should move from polynomials to vectors, and then just work with vectors. It's not the only way, but it seems the easiest. How do you move from polynomials to vectors? Just send a polynomial $p(t)$ to the vector $[p(t_0), \ldots, p(t_n)]$ of its values at the given time points (here, it is $t = -5, -3, -1, 1, 3, 5$). So our polynomials $p_0(t), p_1(t)$ and $p_2(t)$ become vectors

$$p_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \ p_1 = \begin{bmatrix} -5 \\ -3 \\ -1 \\ 1 \\ 3 \\ 5 \end{bmatrix}, \ \text{and} \ \begin{bmatrix} 5 \\ -1 \\ -4 \\ -4 \\ -1 \\ 5 \end{bmatrix}$$

Notice that the inner product of polynomials $\langle p, q \rangle := p(t_0)q(t_0) + \cdots + p(t_n)q(t_n)$ then becomes just the usual dot product of the corresponding vectors. It is also obvious that the three vectors are orthogonal to each other.
Then all that quadratic trend thing is nothing but projecting the data vector g to the three vectors, corresponding to the polynomials (thinking of them actually as polynomials). In other words, the "quadratic trend" is nothing but the projection

$$\widehat{g} = \frac{g \cdot p_0}{p_0 \cdot p_0} p_0(t) + \frac{g \cdot p_1}{p_1 \cdot p_1} p_1(t) + \frac{g \cdot p_2}{p_2 \cdot p_2} p_2(t)$$

where g is the data vector $g = [1, 1, 4, 4, 6, 8]^T$. Compute the coefficients: $\frac{g \cdot p_0}{p_0 \cdot p_0} = \frac{1+1+4+4+6+8}{1^2+1^2+1^2+1^2+1^2+1^2} = \frac{24}{6} = 4$, and similar $\frac{g \cdot p_1}{p_1 \cdot p_1} = \frac{5}{7}$ and $\frac{g \cdot p_2}{p_2 \cdot p_2} = \frac{1}{14}$. Thus the quadratic trend is

$$\widehat{g} = 4 + \frac{5}{7}t + \frac{1}{14}\left(\frac{3}{8}t^2 - \frac{35}{8}\right)$$

Guess who is too lazy to finish this computation.

Problem 4. Find the second order Fourier approximation of the function $f(t) = t - 1$.

Solution. All you have to do is to find projection of $f(t) = t - 1$ onto the space spanned by the functions $\{1, \cos t, \sin t, \cos 2t, \sin 2t\}$ (because I asked for the second order approximation. If it was third, then there would be also functions $\cos 3t$ and $\sin 3t$, and so on). The projection is taken using the inner product $\langle f, g\rangle = \frac{1}{2\pi}\int_0^{2\pi} f(t)g(t)dt$. I won't do the calculation, it is just the projection formula again.

Problem 5. Orthogonally diagonalize the matrix $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$.

Solution. The characteristic equation is $\lambda^2 - 6\lambda + 8 = 0$, and so eigenvalues are $\lambda_1 = 2$ and $\lambda_2 = 4$. Corresponding eigenvectors are $v_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $v_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. They are orthogonal, but don't have length 1. So we rescale them to be unit vectors, and the resulting vectors will be the columns of the matrix P s.t. $A = PDP^T$. We get

$$P = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

and so

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}^T$$

Problem 6. Explain why an equation $Ax = b$ has a solution if and only if b is orthogonal to all solutions of the equation $A^T x = 0$.

Solution. We will unravel what on Earth does "b is orthogonal to all solutions of the equation $A^T x = 0$" mean. Solutions to the equation $A^T x = 0$ is the set $\{x \mid \text{columns of } A \text{ are orthogonal to } x\}$. Indeed, $A^T x$ is a vector, and its entries are exactly dot products of columns of A with x (since for two vectors $u, v \in \mathbb{R}^n$, $u \cdot v = u^T v$). So $A^T x = 0$ exactly means that x is orthogonal to all the columns of A. Thus, x is orthogonal to $Col(A)$, i.e. "all solutions of the equation $A^T x = 0$" really means "$Col(A)^{\perp}$". Then "b is orthogonal to all solutions of the equation $A^T x = 0$" means "b is orthogonal to $Col(A)^{\perp}$", i.e. $b \in (Col(A)^{\perp})^{\perp}$. But $(Col(A)^{\perp})^{\perp} = Col(A)$ (two $\perp$ cancel each other). So "b is orthogonal to all solutions of the equation $A^T x = 0$" means really "$b \in Col(A)$". But that is exactly the necessary and sufficient condition when $Ax = b$ has a solution.

Problem 7. For a quadratic form $Q = 2x_1^2 - 4x_1x_2 - x_2^2$ on $\mathbb{R}^2$ make a change of variable $x = Py$ that transforms the form into one without cross-product terms.

Solution. This is exactly the question of orthogonal diagonalization of the matrix

$$A = \begin{bmatrix} 2 & -2 \\ -2 & -1 \end{bmatrix}$$

corresponding to the quadratic form, that is, A is the matrix such that $Q(x) = x^T Ax$. So please, diagonalize orthogonally this matrix A, find P as in Problem 5, and that's going to be your change of variable.

Problem 8. Same question for the quadratic form $Q = x_1^2 - 12x_1x_2 + 8x_1x_3 + 2x_2^2 - 4x_2x_3 - 3x_3^2$. Help: the eigenvalues of the corresponding matrix are $-3, -6, 9$.

Solution. The same thing as above. You just need to orthogonally diagonalize the matrix

$$A = \begin{bmatrix} 1 & -6 & 4 \\ -6 & 2 & -2 \\ 4 & -2 & -3 \end{bmatrix}$$