Math 414 - Solutions for Assignment 1

Al Boggess

Spring 1998

3. By writing out the inner product, we have

$$\begin{aligned} \langle V, V \rangle &= |v_1|^2 + 4|v_2|^2 + 4Re\{v_1 \overline{v_2}\} \\ &= |v_1 + 2v_2|^2 \end{aligned}$$

So if $V = (v_1, v_2)$ with $v_1 + 2v_2 = 0$, then $\langle V, V \rangle = 0$. This does not define an inner product since there are nonzero vectors (e.g. $V = (-2, 1)$) with $\langle V, V \rangle = 0$.

4. a) (Outline) To show conjugate-symmetry, we have

$$\langle f, g \rangle = \int_a^b f\overline{g} = \overline{\int_a^b g\overline{f}} = \overline{\langle g, f \rangle}$$

To show homogeneous:

$$\langle cf, g \rangle = \int_a^b cf\overline{g} = c\int_a^b f\overline{g} = c\langle f, g \rangle$$

To show bilinearity:

$$\langle f, g + h \rangle = \int_a^b f(\overline{g + h}) = \int_a^b f\overline{g} + \int_a^b f\overline{h} = \langle f, g \rangle + \langle f, h \rangle$$

b) To show positivity for continuous functions, suppose $\int_a^b |f(t)|^2 = 0$ but $|f(t_0)| > 0$ for some t_0. We want to derive a contradiction. The definition of continuity states for each $\epsilon > 0$, there is a $\delta > 0$ such that

$$\text{if } |t - t_0| < \delta \text{ then } |f(t) - f(t_0)| < \epsilon \qquad **$$

Choose $\epsilon = |f(t_0)|/2$. This ϵ is positive since $|f(t_0)|$ is assumed to be positive. $**$ now implies that

$$\text{if } |t - t_0| < \delta \text{ then } |f(t) - f(t_0)| < |f(t_0)|/2$$

which in turn implies

$$|f(t)| > |f(t_0)|/2 \quad \text{for} \quad t_0 - \delta \le t \le t_0 + \delta$$

Integrating this inequality gives

$$\int_a^b |f(t)|^2\,dt > \int_{t_0-\delta}^{t_0+\delta} |f(t_0)|^2/4\,dt = (2\delta)|f(t_0)|^2/4 > 0$$

This contradicts the fact that $\int_a^b |f|^2 = 0$.

5. The conjugate symmetry and bilinearity properties follow easily from manipulating infinite sums. For example

$$\begin{aligned}
\langle x, y\rangle &= \sum_n x_n \overline{y_n} \\
&= \overline{\sum_n y_n \overline{x_n}} \\
&= \overline{\langle y, x\rangle}
\end{aligned}$$

The positivity also follows easily

$$\begin{aligned}
\langle x, x\rangle &= \sum_n x_n \overline{x_n} \\
&= \sum_n |x_n|^2 \ge 0
\end{aligned}$$

The only way the above sum could equal 0 is if each $x_n = 0$ which would mean $x = 0$.

6. We have

$$\int_0^1 |f_n(t)|^2 = \int_0^{1/n} 1\,dt = 1/n \mapsto 0 \quad \text{as } n \mapsto \infty$$

However, f_n does not converge to zero uniformly on [0,1]. To see this, choose $\epsilon = 1/2$ in the definition of uniform convergence. If $f_n \mapsto 0$ uniformly, then there would exist an N so that if $n > N$ then $|f_n(t)| < 1/2$ for all $0 \le t \le 1$. However, $f_n(t) = 1 > 1/2$ for $0 \le t \le 1/n$.

7. Clearly $f_n(0)$ does not converge to zero since $f_n(0) = \sqrt{n} \mapsto \infty$ as $n \mapsto \infty$. Since $f_n(t) = \sqrt{n}$ for $0 \le t \le 1/n^2$ (and is zero elsewhere), we have

$$\int_0^1 |f_n(t)|^2\,dt = \int_0^{1/n^2} (\sqrt{n})^2 = \left(\frac{1}{n^2}\right) n = 1/n$$

which converges to zero as $n \mapsto \infty$.

8. No. For example, $f_n(t) = 1/n$ converges to zero uniformly (easy to show). However $\int_0^\infty |f_n(t)|^2\,dt = \int_0^\infty 1/n^2\,dt = \infty$ for each value of n. Therefore certainly f_n does *not* converge to zero in $L^2[0,\infty]$.

9. The orthogonal complement of the vector $(1,-2,1)$ is the set of all vectors (x,y,z) with

$$\langle (x,y,z),(1,-2,1)\rangle = 0$$

or

$$x - 2y + z = 0$$

(which is the equation of a plane in R^3).

10. If $g(t)$ belongs to the orthogonal complement of $f(t) = 1$ then

$$0 = \langle g, f\rangle = \int_0^1 g(t)\cdot 1\,dt = \int_0^1 g(t)\,dt$$

This is the same as saying that the average value of g is zero.

11. To show that f' is orthogonal to sin in $L^2[0,\pi]$, we must show $\int_0^\pi f'(t)\sin(t)\,dt = 0$. Using integration by parts, we have

$$\begin{aligned}\int_0^\pi f'(t)\sin(t)\,dt &= f(t)\sin(t)|_0^\pi - \int_0^\pi f(t)\cos(t)\,dt \\ &= -\int_0^\pi f(t)\cos(t)\,dt \quad \text{since } \sin(0) = \sin(\pi) = 0 \\ &= -\langle f, \cos\rangle \\ &= 0\end{aligned}$$

by hypothesis. We therefore conclude that $\langle f', \sin\rangle = \int_0^\pi f'(t)\sin(t)\,dt = 0$, as desired.

12. An orthonormal basis for the space spanned by $\{1, x, x^2, x^3\}$ in $L^2[0,1]$ derived from Graham Schmidt is $e_1 = 1$, $e_2 = \sqrt{3}(2x-1)$, $e_3 = 6\sqrt{5}(x^2 - x + 1/6)$, $e_4 = 20\sqrt{7}(x^3 - 3x^2/2 + 3x/5 - 1/20)$.

13. Let f_n be the projection of $f(x) = x^2$ onto the space V_n. The expressions for f_n found from Theorem 4 are the following:

$$\begin{aligned}f_1 &= \pi^3/3 - 4\cos x \\ f_2 &= \pi^3/3 - 4\cos x + \cos 2x \\ f_3 &= \pi^3/3 - 4\cos x + \cos 2x - 4/9\cos 3x\end{aligned}$$

The corresponding f_n for $f(x) = x^3$ are:

$$\begin{aligned}
f_1 &= (2\pi^2 - 12)\sin x \\
f_2 &= (2\pi^2 - 12)\sin x + (3/2 - \pi^2)\sin 2x \\
f_3 &= (2\pi^2 - 12)\sin x + (3/2 - \pi^2)\sin 2x + (2\pi^2/3 - 4/9)\sin 3x
\end{aligned}$$

Note that x^2 is an even function, which is why only cosine terms appear in its projection. Likewise, x^3 is an odd function which is why only sin terms appear in its projection.

14. First note that all four functions $\phi(x)$, $\psi(x)$, $\psi(2x)$ and $\psi(2x-1)$ are orthogonal in L^2 but they do not all have unit length. In fact, the lengths of $\phi(x)$ and $\psi(x)$ are both 1, but the length of $\psi(2x)$ and $\psi(2x-1)$ are both $1/\sqrt{2}$. Therefore the following set of functions form an orthonormal set: $\phi(x)$, $\psi(x)$, $\sqrt{2}\psi(2x)$ and $\sqrt{2}\psi(2x-1)$. The projection of $f(x) = x$ onto the space spanned by these functions is given by

$$\begin{aligned}
< x, \phi > \phi &+ < x, \psi > \psi + < x, \sqrt{2}\psi(2x) > \sqrt{2}\psi(2x) \\
&+ < x, \sqrt{2}\psi(2x-1) > \sqrt{2}\psi(2x-1)
\end{aligned}$$

which evaluates to:

$$\phi(x) - \psi(x)/4 - \psi(2x)/8 - \psi(2x-1)/8$$

15. Suppose $\langle u_0, v\rangle = \langle u_1, v\rangle$, then $\langle u_0 - u_1, v\rangle = 0$ for all v. Letting $v = u_0 - u_1$, we obtain

$$0 = \langle u_0 - u_1, u_0 - u_1\rangle = ||u_0 - u_1||^2$$

which means that $u_0 - u_1 = 0$ or $u_0 = u_1$, as desired.

16. We have

$$\begin{aligned}
\langle T(f), g\rangle &= \langle \int_y K(x,y) f(y)\, dy, g(x)\rangle_x \\
&= \int_{y,x} K(x,y) f(y) \overline{g(x)}\, dy\, dx \\
&= \int_{y,x} f(y) \overline{\overline{K(x,y)} g(x)}\, dy\, dx \\
&= \langle f(y), \int_x \overline{K(x,y)} g(x)\, dx\rangle_y
\end{aligned}$$

On the other hand, $\langle T(f), g\rangle = \langle f, T^*g\rangle$, by definition. Comparing this with the previous equation, we conclude

$$\langle f, T^*g\rangle = \langle f(y), \int_x \overline{K(x,y)}g(x)\,dx\rangle_y$$

which means that $(T^*g)(y) = \int_x \overline{K(x,y)}g(x)\,dx$. By switching the roles of x and y, we obtain

$$T^*g(x) = \int_y \overline{K(y,x)}g(y)\,dy$$

as desired.

17. If $w \in Ker(A^*)$, then $A * w = 0$. For any $v \in V$, we have

$$0 = \langle v, A^*w\rangle = \langle Av, w\rangle$$

and so w is orthogonal to the range of A. Conversely, if w is orthogonal to the range of A, then

$$0 = \langle Av, w\rangle = \langle v, A^*w\rangle \quad \text{for all } v \in V$$

In particular, letting $v = A^*w$, we conclude that $0 = ||A^*w||^2$, which implies $A^*w = 0$.

18. If $Ax = b$ does not have a solution, then b is not in the Range of A. Let w be the projection of b onto $(Range(A))^\perp$. Since b does not belong to $Range(A)$, we have $\langle b, w\rangle \neq 0$. In addition, $A^*w = 0$ by problem 17.

19. For the inclusion $\subset$, suppose $v_0 \in V_0$, then $\langle v_0, w\rangle = 0$ for all $w \in V_0^\perp$. This means that v_0 is orthogonal to everything in $V_0^\perp$ and therefore v_0 belongs to $V_0^{\perp\perp}$.

For the reverse inclusion, suppose w belongs to $V_0^{\perp\perp}$. Decompose w as

$$w = v_0 + v_1 \quad \text{where } \; v_0 \in V_0 \;\; v_1 \in V_0^\perp$$

Since $w \in V_0^{\perp\perp}$, $\langle w, v_1\rangle = 0$. So we have

$$0 = \langle w, v_1\rangle = \langle v_0 + v_1, v_1\rangle = \langle v_1, v_1\rangle$$

where the last equality follows from the fact that $\langle v_0, v_1\rangle = 0$ (since v_1 belongs to $V_0^\perp$). Therefore $||v_1||^2 = 0$ and so $v_1 = 0$. this means $w = v_0$ which belongs to V_0 as desired.

20. Suppose $\{e_j\}_{j=0}^{\infty}$ is orthonormal. To show that this set is linearly independent, we must show that if

$$\sum_j \alpha_j e_j = 0$$

then each $\alpha_j = 0$. Taking the inner product of the above sum with e_k we get

$$0 = \langle \sum_j \alpha_j e_j, e_k \rangle = \alpha_k \langle e_k, e_k \rangle$$

where the last equality uses $\langle e_j, e_k \rangle = 0$ if $j \neq k$. Since $\langle e_k, e_k \rangle = 1$, the right side is just α_k. Therefore, we conclude that $\alpha_k = 0$ as desired.

21. The matrix equation

$$Z^t Z = \begin{pmatrix} m \\ b \end{pmatrix} = Z^t Y$$

is equivalent to

$$\begin{pmatrix} |x|^2 & N\overline{x} \\ N\overline{x} & N \end{pmatrix} \begin{pmatrix} m \\ b \end{pmatrix} = \begin{pmatrix} x \cdot y \\ N\overline{y} \end{pmatrix} \tag{1}$$

The inverse of the matrix on the left is

$$\frac{1}{\sigma} \begin{pmatrix} 1 & -\overline{x} \\ -\overline{x} & |x|^2/N \end{pmatrix}$$

Multiplying this matrix on both sides of the above equation gives the desired result for m and b. Note that there are two equivalent expressions for σ. This is just the variance (statistics) of the x_i, which is

$$\begin{aligned} \sigma &= \sum_i (x_i - \overline{x})^2 \\ &= \sum_{i=1}^{N} (x_i^2 - 2x_i\overline{x} + \overline{x}^2) \\ &= |x|^2 - 2N\overline{x}^2 + N\overline{x}^2 \end{aligned}$$

where the last equation uses the notation $\overline{x} = \frac{1}{N}\sum_i x_i$ and the fact that summing $\overline{x}$ N-times, yields $N\overline{x}^2$. This last expression is just $|x|^2 - N\overline{x}^2$, which is the second expression for σ.

22. Following the outline of the derivation of Theorem 6, we try to minimize the quantity

$$E = \sum_{i=1}^{N} |y_i - (ax_i^2 + bx_i + c)|^2$$

The quantity E can be viewed as the square of the distance (in R^N) from the vector

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix}$$

and the vector $aX_2 + bX_1 + cU$ where

$$X_2 = \begin{pmatrix} x_1^2 \\ x_2^2 \\ \vdots \\ x_N^2 \end{pmatrix} \qquad X_1 = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} \qquad U = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

As a, b, and c vary over all possible real numbers the expression $aX_2 + bX_1 + U$ sweeps out a three dimensional vector subspace of R^N which is generated by X_2, X_1 and U. Thus we are interested in finding the point $P = aX_2 + bX_1 + cU$ on this subspace which is closest to the point Y. P must be chosen so that $Y - P$ is orthogonal to this subspace, which means that $Y - P$ must be orthogonal to the vectors X_2, X_1 and U. These orthogonality relations can be written in matrix form as

$$\begin{pmatrix} x_1^2 & \dots & x_N^2 \\ x_1 & \dots & x_N \\ 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix} = \begin{pmatrix} x_1^2 & \dots & x_N^2 \\ x_1 & \dots & x_N \\ 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} x_1^2 & x_1 & 1 \\ \vdots & \vdots & \vdots \\ x_N^2 & x_N & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

This matrix equation can be written in shorthand form as

$$Z^t Y = Z^t Z \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

where

$$Z = \begin{pmatrix} x_1^2 & x_1 & 1 \\ \vdots & \vdots & \vdots \\ x_N^2 & x_N & 1 \end{pmatrix}$$

as desired.

23. Set

$$E = \sum_{i=1}^{N} (mx_i + b - y_i)^2$$

E will be minimized when

$$\begin{aligned} 0 &= \frac{\partial E}{\partial m} = 2\sum_{i=1}^{N}(mx_i + b - y_i)x_i \\ 0 &= \frac{\partial E}{\partial b} = 2\sum_{i=1}^{N}(mx_i + b - y_i) \end{aligned}$$

These equations can be rewritten using $\overline{x} = (1/N)\sum_i x_i$ as

$$\begin{aligned} m|x|^2 + bN\overline{x} &= x \cdot y \\ mN\overline{x} + Nb &= N\overline{y} \end{aligned}$$

or in matrix terms

$$\begin{pmatrix} |x|^2 & N\overline{x} \\ N\overline{x} & N \end{pmatrix} \begin{pmatrix} m \\ b \end{pmatrix} = \begin{pmatrix} x \cdot y \\ N\overline{y} \end{pmatrix}$$

which is the same matrix equation as in equation (1). Therefore the calculus derivation leads to the same equation as the one presented in the text and thus has the same solution for m and b.

24. The answer for the best fit line is $y = 4x + 1$.

25. The answer for the best fit quadratic is $y = 2 + 4x/3 + 2x^2/3$.