**MATH 725 Homework Set 8 due 10/17/11**

(1) Let $U = \operatorname{span}((1,1,1,0),(1,1,0,2)) \subseteq \mathbf{R}^4$. Find the vector $\mathbf{u} \in U$ that minimizes $||\mathbf{u} - (1,2,3,4)||$.

*Solution.* We consider $\mathbf{R}^4$ with the usual inner product and find an orthnormal basis of $U$: use Gram–Schmidt on $((1,1,1,0),(1,1,0,2))$ to obtain the orthonormal basis $(\mathbf{e}_1, \mathbf{e}_2)$ where

$$\mathbf{e}_1 := \tfrac{1}{\sqrt{3}}(1,1,1,0)$$

and

$$\mathbf{e}_2 := \frac{(1,1,0,2) - \langle (1,1,0,2), \mathbf{e}_1 \rangle \mathbf{e}_1}{||(1,1,0,2) - \langle (1,1,0,2), \mathbf{e}_1 \rangle \mathbf{e}_1||} = \sqrt{\tfrac{3}{14}} \left( \tfrac{1}{3}, \tfrac{1}{3}, -\tfrac{2}{3}, 2 \right).$$

We proved in class that the projection of $\mathbf{v} \in \mathbf{R}^4$ onto $U$ is then given by

$$\mathbf{u} := \langle \mathbf{v}, \mathbf{e}_1 \rangle \mathbf{e}_1 + \langle \mathbf{v}, \mathbf{e}_2 \rangle \mathbf{e}_2$$

and this $\mathbf{u}$ minimizes the distance to $\mathbf{v}$ amond all vectors in $U$. For $\mathbf{v} = (1,2,3,4)$, this is

$$\begin{aligned}\mathbf{u} &= \left\langle (1,2,3,4), \tfrac{1}{\sqrt{3}}(1,1,1,0) \right\rangle \tfrac{1}{\sqrt{3}}(1,1,1,0) + \left\langle (1,2,3,4), \sqrt{\tfrac{3}{14}} \left( \tfrac{1}{3}, \tfrac{1}{3}, -\tfrac{2}{3}, 2 \right) \right\rangle \sqrt{\tfrac{3}{14}} \left( \tfrac{1}{3}, \tfrac{1}{3}, -\tfrac{2}{3}, 2 \right) \\ &= \left( \tfrac{5}{2}, \tfrac{5}{2}, 1, 3 \right)\end{aligned}$$

□

(2) Find the polynomial $p(x) \in \mathscr{P}_3(\mathbf{R})$ with zero constant and linear coefficient that minimizes

$$\int_{-1}^{1} |2 + 3x - p(x)|^2 \, dx.$$

*Solution.* We use our inner product from last week:

$$\langle f, g \rangle := \int_{-1}^{1} f(x) \, g(x) \, dx.$$

Let $U := \left\{ ax^3 + bx^2 : \, a, b \in \mathbf{R} \right\}$ and $q(x) := 2 + 3x$; then the problem asks for the closest point $p \in U$ to $q$, i.e., the orthogonal projection of $q$ onto $U$.

We first compute an orthonormal basis of $U$. A basis of $U$ is $\left( x^2, x^3 \right)$; applying Gram–Schmidt this basis gives

$$e_1(x) = \frac{x^2}{||x^2||} = \sqrt{\tfrac{5}{2}} \, x^2$$

and

$$e_2(x) = \frac{x^3 - \langle x^3, e_1(x) \rangle e_1(x)}{||x^3 - \langle x^3, e_1(x) \rangle e_1(x)||} = \frac{x^3}{||x^3||} = \sqrt{\tfrac{7}{2}} \, x^3.$$

We know that the projection of $q$ onto $U$ is

$$p = \langle q, e_1 \rangle e_1 + \langle q, e_2 \rangle e_2 = \tfrac{5}{2} x^2 \int_{-1}^{1} (2+3x) x^2 \, dx + \tfrac{7}{2} x^3 \int_{-1}^{1} (2+3x) x^3 \, dx = \tfrac{10}{3} x^2 + \tfrac{21}{5} x^3.$$

□

(3) Suppose $T \in L(V)$. Prove that if $T^2 = T$ and every vector in $\operatorname{null}(T)$ is orthogonal to every vector in $\operatorname{range}(T)$, then $T$ is an orthogonal projection onto $\operatorname{range}(T)$.

*Proof.* Suppose that $T^2 = T$ and every vector in $\operatorname{null}(T)$ is orthogonal to every vector in $\operatorname{range}(T)$. Then we can write any $\mathbf{v} \in V$ as

$$\mathbf{v} = T(\mathbf{v}) + (\mathbf{v} - T(\mathbf{v})).$$

The first summand on the right-hand side is in $\operatorname{range}(T)$, and so we're done if we can prove that the second summand, $\mathbf{v} - T(\mathbf{v})$, is in $\operatorname{range}(T)^\perp$. But this follows from

$$T(\mathbf{v} - T(\mathbf{v})) = T(\mathbf{v}) - T^2(\mathbf{v}) = 0,$$

which means that $\mathbf{v} - T(\mathbf{v})$ is in $\operatorname{null}(T)$ and thus orthogonal to every vector in $\operatorname{range}(T)$. □

(4) Suppose $T \in L(V)$.
  (a) Prove that $\lambda$ is an eigenvalue of $T$ if and only if $\overline{\lambda}$ is an eigenvalue of $T^*$.
  (b) Show that $U$ is invariant under $T$ if and only if $U^\perp$ is invariant under $T^*$.

*Proof.* (a) We will prove the contrapositive (in both directions). $\lambda$ is not an eigenvalue of $T$ if and only if $T - \lambda\mathrm{I}$ is invertible, which means there exists $S \in L(V)$ such that

$$S(T - \lambda\mathrm{I}) = (T - \lambda\mathrm{I})S = \mathrm{I}.$$

This, in turn, is equivalent to

$$(T - \lambda\mathrm{I})^* S^* = S^*(T - \lambda\mathrm{I})^* = \mathrm{I},$$

which means that $(T - \lambda\mathrm{I})^* = T^* - \overline{\lambda}\mathrm{I}$ is invertible, which holds if and only if $\overline{\lambda}$ is not an eigenvalue of $T^*$.

(b) We only have to prove one implication, since $(U^\perp)^\perp = U$ and $(T^*)^* = T$. Suppose $U$ is invariant under $T$. Given $\mathbf{v} \in U^\perp$, we need to show that $T^*(\mathbf{v}) \in U^\perp$, i.e., that $\langle \mathbf{u}, T^*(\mathbf{v})\rangle = 0$ for all $\mathbf{u} \in U$. But for any such $\mathbf{u} \in U$,

$$\langle \mathbf{u}, T^*(\mathbf{v})\rangle = \langle T(\mathbf{u}), \mathbf{v}\rangle = 0$$

because $T(\mathbf{u}) \in U$ and $\mathbf{v} \in U^\perp$. □

(5) Suppose $A \in L(\mathbf{R}^n, \mathbf{R}^m)$ is given by a matrix (written in terms of the standard bases) whose columns are linearly independent, and let $\mathbf{b} \in \mathbf{R}^m$. If $A$ is not surjective, the linear system $A\mathbf{x} = \mathbf{b}$ might not have a solution $\mathbf{x} \in \mathbf{R}^n$. The following exercise computes the best approximation (in the sense of part (c)) to a solution.
  (a) Show that $A$ is injective.
  (b) Show that $A^*A$ is invertible.
  (c) Prove that $\mathbf{y} := (A^*A)^{-1}A^*\mathbf{b}$ satisfies $||A\mathbf{y} - \mathbf{b}|| \le ||A\mathbf{x} - \mathbf{b}||$ for all $\mathbf{x} \in \mathbf{R}^n$.[1]

*Proof.* (a) Suppose the columns of $A$ are $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n \in \mathbf{R}^m$. A vector $(v_1, v_2, \ldots, v_n) \in \operatorname{null}(A)$ satisfies

$$v_1\mathbf{c}_1 + v_2\mathbf{c}_2 + \ldots v_n\mathbf{c}_n = \mathbf{0},$$

and so $v_1 = v_2 = \cdots = v_n = 0$ because $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n$ are linearly independent. Thus $\operatorname{null}(A) = \{\mathbf{0}\}$.

(b) Since $A^*A \in L(\mathbf{R}^n)$, so it suffices to prove that $\operatorname{null}(A^*A) = \{\mathbf{0}\}$. Suppose $A^*A\mathbf{v} = \mathbf{0}$. Then

$$0 = \langle \mathbf{v}, A^*A\mathbf{v}\rangle = \langle A\mathbf{v}, A\mathbf{v}\rangle$$

and so $A(\mathbf{v}) = \mathbf{0}$. Part (a) implies that $\mathbf{v} = \mathbf{0}$. Thus $\operatorname{null}(A^*A) = \{\mathbf{0}\}$.

(c) We claim that $P := A(A^*A)^{-1}A^*$ is the orthogonal projection onto $\operatorname{range}(A)$.
Note that $A^*$ is surjective (because $A$ is injective), as is $(A^*A)^{-1}$ (because it is invertible), and so $\operatorname{range}(P) = \operatorname{range}(A)$. Thus, by Exercise (3), it suffices to prove that $P^2 = P$ and that every vector in $\operatorname{null}(P)$ is orthogonal to every vector in $\operatorname{range}(A)$.
The former assertion follows with $P^2 = A\left((A^*A)^{-1}A^*A\right)(A^*A)^{-1}A^* = A(A^*A)^{-1}A^* = P$.
To prove that every vector in $\operatorname{null}(P)$ is orthogonal to every vector in $\operatorname{range}(A)$, suppose $\mathbf{v} \in \operatorname{null}(P)$. Then $A(A^*A)^{-1}A^*\mathbf{v} = \mathbf{0}$, i.e., $(A^*A)^{-1}A^*\mathbf{v} \in \operatorname{null}(A)$, which implies by (a) that $(A^*A)^{-1}A^*\mathbf{v} = \mathbf{0}$, which implies by (b) that $A^*\mathbf{v} = \mathbf{0}$, i.e., $\mathbf{v} \in \operatorname{null}(A^*)$. We proved in class that $\operatorname{null}(A^*) = \operatorname{range}(A)^\perp$, and so $\mathbf{v}$ is orthogonal to every vector in $\operatorname{range}(A)$.
Thus we have proved that $P$ is the orthogonal projection onto $\operatorname{range}(A)$. But this means, given $\mathbf{b} \in \mathbf{R}^m$, $P\mathbf{b}$ is closest (measured using the norm) to $\mathbf{b}$ among all the vectors in $\operatorname{range}(A)$, in other words,

$$||A\mathbf{y} - \mathbf{b}|| = ||P\mathbf{b} - \mathbf{b}|| \le ||A\mathbf{x} - \mathbf{b}||$$

for all $\mathbf{x} \in \mathbf{R}^n$. □

[1] *Hint:* Start by proving that $A(A^*A)^{-1}A^*$ is the orthogonal projection onto $\operatorname{range}(A)$.