Selected Solutions to Homwork 6

Tim Smits

5.2.27 Find the QR decomposition of $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 2 \\ 1 & 0 & 1 \\ 1 & 1 & -1 \end{pmatrix}$.

Solution: Let the columns of A be v_1, v_2, v_3. Set $u_1 = v_1/\|v_1\| = v_1/2 = (1/2, 1/2, 1/2, 1/2)$. Then $v_2^\perp = v_2 - (u_1 \cdot v_2)u_1 = v_2 - u_1 = (1/2, -1/2, -1/2, 1/2)$. Since $v_2^\perp$ is a unit vector, $v_2 = u_2$. We then take $v_3^\perp = v_3 - (u_1 \cdot v_3)u_1 - (u_2 \cdot v_3)u_2 = v_3 - u_1 + 2u_2 = (1/2, 1/2, -1/2, -1/2)$, which is a unit vector, so $u_3 = v_3^\perp$. We then have $v_1 = 2u_1$, $v_2 = u_1 + u_2$ and $v_3 = u_1 - 2u_2 + u_3$, so $Q = \begin{pmatrix} 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 \\ 1/2 & 1/2 & -1/2 \end{pmatrix}$ and $R = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$.

5.3.27 Let A be an $n \times m$ matrix, $v \in \mathbb{R}^m$ and $w \in \mathbb{R}^n$. Show that $(Av) \cdot w = v \cdot (A^t w)$.

Solution: We have $(Av) \cdot w = w^t Av$. On the other hand, $v \cdot (A^t w) = (A^t w)^t v = w^t Av$.

5.3.29 Show that an orthogonal linear transformation $L : \mathbb{R}^n \to \mathbb{R}^n$ preserves angles: the angle between v and w in $\mathbb{R}^n$ is the same as that of $L(v)$ and $L(w)$. Conversely, is any linear transformation that preserves angles orthogonal?

Solution: Let A be the matrix of L. Since L is orthogonal, $L(v) \cdot L(w) = (Av) \cdot (Aw) = w^t A^t Av = w^t v = v \cdot w$, since A is an orthogonal matrix. We have $v \cdot w = \|v\|\|w\| \cos(\theta)$ where θ is the angle between v and w. Similarly, $L(v) \cdot L(w) = \|L(v)\|\|L(w)\| \cos(\theta')$ where θ' is the angle between $L(v)$ and $L(w)$. Since L is orthogonal, we have $\|L(v)\| = \|v\|$ and $\|L(w)\| = \|w\|$, so this then says $\cos(\theta) = \cos(\theta')$. Since $0 \leq \theta, \theta' \leq \pi$, this then says $\theta = \theta'$.

The converse is false; take for example, $L(x) = 2x$, which just scales vectors by a factor of two. Then L is not orthogonal because it doesn't preserve length, but it obviously preserves angles.

5.3.31 Are the rows of an orthogonal matrix orthonormal?

Solution: Yes – since A is orthogonal, $A^{-1} = A^t$, which then tells us that A^t is also orthogonal. The columns of A^t are then orthonormal, but these are of course just the rows of A.

5.3.69 Find the matrix of Proj_W where $W = \text{Span}\{(1, 1, -1, 0), (0, 1, 1, -1)\}$.

Solution: An orthonormal basis of W is given by $\{(1/\sqrt{3}, 1/\sqrt{3}, -1/\sqrt{3}, 0), (0, 1/\sqrt{3}, 1/\sqrt{3}, -1/\sqrt{3})\}$, since the given vectors are already orthogonal. Let $Q = \begin{pmatrix} 1/\sqrt{3} & 0 \\ 1/\sqrt{3} & 1/\sqrt{3} \\ -1/\sqrt{3} & 1/\sqrt{3} \\ 0 & -1/\sqrt{3} \end{pmatrix}$. Then $\text{Proj}_W =$ $QQ^t = \begin{pmatrix} 1/3 & 1/3 & -1/3 & 0 \\ 1/3 & 2/3 & 0 & -1/3 \\ -1/3 & 0 & 2/3 & -1/3 \\ 0 & -1/3 & -1/3 & 1/3 \end{pmatrix}$.

5.4.3 Let V be a subspace of $\mathbb{R}^n$. Let $v_1, \ldots, v_p$ be a basis of V and $w_1, \ldots, w_q$ be a basis of $V^\perp$. Are $v_1, \ldots, v_p, w_1, \ldots, w_q$ a basis of $\mathbb{R}^n$?

Solution: Yes – since $p+q = \dim(V)+\dim(V^\perp) = n$, we just need to show that $v_1, \ldots, v_p, w_1, \ldots, w_q$ are linearly independent. Suppose that $c_1v_1 + \ldots + c_pv_p + d_1w_1 + \ldots + d_qw_q = 0$. Then subtracting says $c_1v_1 + \ldots + c_pv_p = -d_1w_1 - \ldots - d_qw_q$. The right hand side is a vector in $V^\perp$, while the left hand side is a vector in V. Since $V \cap V^\perp = \{0\}$, this says $c_1v_1 + \ldots + c_pv_p = 0$. Since the v_i are a basis of V, we see all $c_i = 0$. From our original equation, we then have $d_1w_1 + \ldots + d_qw_q = 0$, and since w_i are a basis of $V^\perp$, this says all $d_i = 0$. This shows $v_1, \ldots, v_p, w_1, \ldots, w_q$ are linearly independent vectors, and therefore a basis of $\mathbb{R}^n$.

5.4.11 Consider $L(x) = Ax$, a linear transformation from $\mathbb{R}^n \to \mathbb{R}^m$ where A is an $m \times n$ matrix of rank m. Define the *pseudo-inverse* of L, L^+, to be the linear transformation $L^+ : \mathbb{R}^m \to \mathbb{R}^n$ with $L^+(y) = \{\text{the minimal solution to } L(x) = y\}$.

(a) Show that L^+ is linear.

(b) What is $L(L^+(y))$?

(c) What is $L^+(L(x))$?

(d) Determine $\text{im}(L^+)$ and $\ker(L^+)$.

(e) Find L^+ for $L(x) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} x$.

Solution:

(a) From problem 5.4.10, the minimal solution of $L(x) = y$ is defined as x_0, where $Ax = y$, $x = x_h + x_0$, for $x_h \in \ker(L)$ and $x_0 \in \ker(L)^\perp$. That is, $x_0 = \text{Proj}_{\ker(L)^\perp}(x)$. Unraveling the definition here, we have $Ax = Ax_0 = y$, and $L^+(y) = x_0$. How can we solve for x_0 in terms of y? Since $\ker(L)^\perp = \text{im}(L^t)$, we have $x_0 = \text{Proj}_{\text{im}(L^t)}(x)$, so $x_0 = A^t x^*$ where x^* is the least squares solution to the system $A^t x^* = x$ (make sure this is clear, this is the key step!). We then have $x^* = (AA^t)^{-1}Ax = (AA^t)^{-1}y$, so $x_0 = A^t x^* = A^t(AA^t)^{-1}y$. This says $L^+(y) = A^t(AA^t)^{-1}y$, so in particular, L^+ is just multiplication by some matrix and therefore is a linear transformation.

(b) We have $L(L^+(y)) = L(A^t(AA^t)^{-1}y) = A(A^t(AA^t)^{-1}y) = y$.

(c) We have $L^+(L(x)) = L^+(Ax) = A^t(AA^t)^{-1}Ax$.

(d) I claim that $\text{im}(L^+) = \text{im}(L^t)$ and $\ker(L^+) = \ker(L^t)$. If $y \in \text{im}(L^t)$, then $y = A^t(AA^t)^{-1}x$ for some x, so $y = A^t((AA^t)^{-1}x) \in \text{im}(L^t)$ says $\text{im}(L^+) \subset \text{im}(L^t)$. If $y \in \text{im}(L^t)$, then $y = A^t x$ for some x. Then $y = A^t x = A^t(AA^t)^{-1}(AA^t x) = L^+(AA^t x)$, so $y \in \text{im}(L^+)$ gives $\text{im}(L^t) = \text{im}(L^+)$. The proof that $\ker(L^t) = \ker(L^+)$ is similar.

(e) $L^+(x) = A^t(AA^t)^{-1}x = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} x.$

5.4.13 Now let $L(x) = Ax$ be a linear transformation from $\mathbb{R}^n \to \mathbb{R}^m$ (now with no conditions on $\text{rank}(A)$). Define the *psuedo-inverse* of L^+ to be $L^+(y) = \{$the minimal least squares solution to $L(x) = y\}$.

(a) Show that L^+ is linear.

(b) What is $L^+(L(x))$?

(c) What is $L(L^+(y))$?

(d) Determine $\text{im}(L^+)$ and $\ker(L^+)$.

(e) Find L^+ for $L(x) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} x$.

Solution:

(a) Once more, the hardest part of this question is figuring out what everything means. The minimal least squares solution to $Ax = y$ is given by x_0, where $x^* = x_h + x_0$ for $x_h \in \ker(A)$ and $x_0 \in \ker(A)^\perp$, and x^* is a least squares solution to $Ax = y$ (note: x_0 is independent of the choice of least squares solution by problem 10). Another way of saying this is that $x_0 = \text{Proj}_{\ker(A)^\perp}(x^*)$. Notice that the definition is analogous to before, but we must work with x^* (since y might not be in $\text{im}(A)$!).

Unlike before, we won't be able to get a nice expression for x_0 in terms of y, so we must show L^+ is linear directly. Suppose $L^+(y) = x_0$ and $L^+(y') = x_0'$. Then if x^* a least square solution to $Ax = y$ and $x^{*\prime}$ is a least squares solution to $Ax = y'$, then $A^tA(x^* + x^{*\prime}) = A^tAx^* + A^tAx^{*\prime} = A^ty + A^ty' = A^t(y + y')$, so $x^* + x^{*\prime}$ is a least squares solution to $Ax = y + y'$. This says $L^+(y + y') = \text{Proj}_{\ker(A)^\perp}(x^* + x^{*\prime}) = \text{Proj}_{\ker(A)^\perp}(x^*) + \text{Proj}_{\ker(A)^\perp}(x^{*\prime}) = L^+(y) + L^+(y')$. Similarly, we can show $L^+(ky) = kL^+(y)$, so that L^+ is linear.

(b) $L^+(L(x)) = L^+(Ax)$. The least squares solution to the system $Ax^* = Ax$ is simply $x^* = x$, so $L^+(Ax) = \text{Proj}_{\text{im}(A^t)}(x^*) = \text{Proj}_{\text{im}(A^t)}(x)$.

(c) $L(L^+(y)) = A(L^+(y))$. Let x^* be the least squares solution to $Ax = y$. Then by definition, $L^+(y)$ is x_0 where $x^* = x_h + x_0$ and $x_h \in \ker(A)$ and $x^* \in \ker(A)^\perp$. This says $A(L^+(y)) = Ax_0 = Ax^* = \text{Proj}_{\text{im}(A)}(y)$.

(d) By definition, we have $\text{im}(L^+) \subset \ker(A)^\perp = \text{im}(A^t)$. Given $x_0 = \text{im}(A^t)$, then by part b) we have $L^+(Ax_0) = \text{Proj}_{\text{im}(A^t)}(x_0) = x_0$, so $\text{im}(A^t) \subset \text{im}(L^+)$ says $\text{im}(L^+) = \text{im}(A^t)$. We also see that $L^+(y) = 0$ if and only if $\text{Proj}_{\text{im}(A^t)}(x^*) = 0$, where x^* is the least squares solution to $Ax = y$. This is equivalent to saying that $x^* \in \text{im}(A^t)^\perp = \ker(A)$, so that $Ax^* = 0$. However, this is equivalent to saying that $\text{Proj}_{\text{im}(A)}(y) = 0$, which happens if and only if $y \in \text{im}(A)^\perp = \ker(A^t)$. This shows $\ker(L^+) = \ker(A^t)$.

(e) To compute the matrix of $L^+ : \mathbb{R}^2 \to \mathbb{R}^3$, we compute $L^+(e_1)$ and $L^+(e_2)$, where e_1 and e_2 are the standard basis vectors of $\mathbb{R}^2$. We have $L^+(e_1) = \text{Proj}_{\text{im}(A^t)}(x^*)$, where x^* is the least squares solution to $Ax = e_1$. We compute that $x^* = (1/2, 0, 0)$, and this is already in $\text{im}(A^t)$, so $x_0 = (1/2, 0, 0)$. We similarly find $L^+(e_2) = (0, 0, 0)$, so $L^+(y) = \begin{pmatrix} 1/2 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} y$.

5.4.19 Find the least squares solution to $Ax = b$ where $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

Solution: We need to solve $A^t A x^* = A^t b$. Since $A^t A = I$, this says $x^* = A^t b = (1, 1)$.

5.4.35 Suppose you have a function $g(t)$ and you want to fit it to the form $f(t) = c + p\sin(t) + q\cos(t)$ on the interval $[0, 2\pi]$. The coefficients c, p, q can be determined using linear algebra as follows. First, pick n points $(a_i, g(a_i))$ (assume equally spaced, so $a_i = a_i = \frac{2\pi}{n} i$). We can find a function $f_n(t) = c_n + p_n \sin(t) + q_n \cos(t)$ that "best fits" these data points by solving the relevant least squares equation. This lets us solve for c_n, p_n, q_n. Taking a limit as $n \to \infty$ will then give c, p, q.

Let $A_n = \begin{pmatrix} 1 & \sin(a_1) & \cos(a_1) \\ 1 & \sin(a_2) & \cos(a_2) \\ \vdots & \vdots & \vdots \\ 1 & \sin(a_n) & \cos(a_n) \end{pmatrix}$ and $b_n = \begin{pmatrix} g(a_1) \\ g(a_2) \\ \vdots \\ g(a_n) \end{pmatrix}$.

(a) Compute $A_n^t A_n$ and $A_n^t b_n$.

(b) Compute $\lim_{n\to\infty} \frac{2\pi}{n} A_n^t A_n$ and $\lim_{n\to\infty} \frac{2\pi}{n} A_n^t b_n$.

(c) Find $\lim_{n\to\infty}(c_n, p_n, q_n) = (\lim_{n\to\infty} \frac{2\pi}{n} A_n^t A_n)^{-1} \lim_{n\to\infty} \frac{2\pi}{n} A_n^t b_n$.

Solution:

(a) The (i, j)-th entry of $A_n^t A_n$ is given by the i-th row of A_n^t dotted by the j-th column of A_n because of how matrix multiplication works. It's then not hard to compute $A_n^t A_n = \begin{pmatrix} n & \sum_{i=1}^n \sin(a_i) & \sum_{i=1}^n \cos(a_i) \\ \sum_{i=1}^n \sin(a_i) & \sum_{i=1}^n \sin^2(a_i) & \sum_{i=1}^n \sin(a_i)\cos(a_i) \\ \sum_{i=1}^n \cos(a_i) & \sum_{i=1}^n \sin(a_i)\cos(a_i) & \sum_{i=1}^n \cos^2(a_i) \end{pmatrix}$. Similarly, the i-th entry of $A_n^t b_n$ is the i-th row of A_n^t dotted with b_n. We find $A_n^t b_n = \begin{pmatrix} \sum_{i=1}^n g(a_i) \\ \sum_{i=1}^n g(a_i)\sin(a_i) \\ \sum_{i=1}^n g(a_i)\cos(a_i) \end{pmatrix}$.

(b) Recall the Riemann sum definition of the integral, that $\int_a^b f(x)\,dx = \lim_{n\to\infty} \sum_{i=1}^n f(a + i\frac{b-a}{n})\frac{b-a}{n}$.
In our case, we are trying to compute expressions of the form $\lim_{n\to\infty} \frac{2\pi}{n} \sum_{i=1}^n f(a_i) = \lim_{n\to\infty} \frac{2\pi}{n} \sum_{i=1}^n f(\frac{2\pi}{n}) = \int_0^{2\pi} f(x)\,dx$.
Therefore, $\lim_{n\to\infty} A_n^t A_n = \begin{pmatrix} \int_0^{2\pi} 1\,dx & \int_0^{2\pi} \sin(x)\,dx & \int_0^{2\pi} \cos(x)\,dx \\ \int_0^{2\pi} \sin(x)\,dx & \int_0^{2\pi} \sin^2(x)\,dx & \int_0^{2\pi} \sin(x)\cos(x)\,dx \\ \int_0^{2\pi} \cos(x)\,dx & \int_0^{2\pi} \sin(x)\cos(x)\,dx & \int_0^{2\pi} \cos^2(x)\,dx \end{pmatrix} = \begin{pmatrix} 2\pi & 0 & 0 \\ 0 & \pi & 0 \\ 0 & 0 & \pi \end{pmatrix}$. Similarly, we find $\lim_{n\to\infty} \frac{2\pi}{n} A_n^t b_n = \begin{pmatrix} \int_0^{2\pi} g(x)\,dx \\ \int_0^{2\pi} g(x)\sin(x)\,dx \\ \int_0^{2\pi} g(x)\cos(x)\,dx \end{pmatrix}$

(c) We have $(\lim_{n\to\infty} A_n^t A_n)^{-1} = \begin{pmatrix} \frac{1}{2\pi} & 0 & 0 \\ 0 & \frac{1}{\pi} & 0 \\ 0 & 0 & \frac{1}{\pi} \end{pmatrix}$, so $(\lim_{n\to\infty} \frac{2\pi}{n} A_n^t A_n)^{-1} \lim_{n\to\infty} \frac{2\pi}{n} A_n^t b = \begin{pmatrix} \frac{1}{2\pi}\int_0^{2\pi} g(x)\,dx \\ \frac{1}{\pi}\int_0^{2\pi} g(x)\sin(x)\,dx \\ \frac{1}{\pi}\int_0^{2\pi} g(x)\cos(x)\,dx \end{pmatrix}$.

The function $c + p\sin(x) + q\cos(x)$ is called the first order *trigonometric polynomial* of $g(x)$. Recall from calculus that a "nice" function $f(x)$ may be written as $f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(c)}{n!}(x-c)^n$, the Taylor series of $f(x)$ expanded around c, and that the N-th partial sums of the Taylor series are called N-th order Taylor polynomials. We have another type of series representation, called a *Fourier series* for "nice" 2π-periodic functions $f(x)$, of the form $f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n\cos(nx) + \sum_{n=1}^{\infty} b_n\sin(nx)$. The N-th partial sum of the Fourier series is called the N-th order *trigonometric polynomial* of $f(x)$. The first order trigonometric polynomial of $f(x)$ is of the form $\frac{a_0}{2} + a_1\cos(x) + b_1\sin(x)$. What we just showed (using linear algebra!) was that $a_0 = \frac{1}{\pi}\int_0^{2\pi} f(x)\,dx$, $a_1 = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos(x)\,dx$ and $b_1 = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin(x)\,dx$. In general, we have $a_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos(nx)\,dx$ and $b_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin(nx)\,dx$. Fourier series have extremely important applications in both pure and applied math, computer science, all disciplines of engineering, and any physical science (or sub-field) you can imagine. For a better idea of how Fourier series are related to linear algebra, see section 5.5 of the textbook.