

MAS212 Linear Algebra I 2007-2008

Coursework 10 — Solutions

Exercise* 1.

- (a) Since Q is a *square* matrix and $Q^T Q = I$, Corollary 2.42 implies that $Q Q^T = I$ as well. Thus Q is invertible with $Q^{-1} = Q^T$.
- (b) If $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is an orthonormal basis for $\mathbb{R}^n$, then

$$\mathbf{v}_i \cdot \mathbf{v}_j = \delta_{ij}.$$

By Exercise 7 (a) on Coursework 9, $(Q\mathbf{x}) \cdot (Q\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$ for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Thus

$$(Q\mathbf{v}_i) \cdot (Q\mathbf{v}_j) = \mathbf{v}_i \cdot \mathbf{v}_j = \delta_{ij},$$

so $\{Q\mathbf{v}_1, \dots, Q\mathbf{v}_n\}$ is an orthonormal set in $\mathbb{R}^n$. In particular, the n vectors $Q\mathbf{v}_1, \dots, Q\mathbf{v}_n$ are linearly independent. But $\dim \mathbb{R}^n = n$, so $\{Q\mathbf{v}_1, \dots, Q\mathbf{v}_n\}$ must be a basis of $\mathbb{R}^n$.

Exercise* 2. By the Orthogonal Decomposition Theorem, $\mathbf{y}$ can be written as $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$ with $\hat{\mathbf{y}} \in S$ and $\mathbf{z}$ orthogonal to S , by letting

$$\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \mathbf{u}_1 + \frac{\mathbf{y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2} \mathbf{u}_2 = \frac{6}{3} \mathbf{u}_1 + \frac{7}{14} \mathbf{u}_2 = 2 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} -1 \\ 3 \\ -2 \end{pmatrix} = \begin{pmatrix} 3/2 \\ 7/2 \\ 1 \end{pmatrix}$$

and

$$\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}} = \begin{pmatrix} -1 \\ 4 \\ 3 \end{pmatrix} - \begin{pmatrix} 3/2 \\ 7/2 \\ 1 \end{pmatrix} = \begin{pmatrix} -5/2 \\ 1/2 \\ 2 \end{pmatrix}.$$

Exercise* 3.

- (a) Since $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ and $\{\mathbf{w}_1, \dots, \mathbf{w}_q\}$ are both orthogonal sets, any two distinct vectors in $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ are orthogonal, and any two distinct vectors in $\{\mathbf{w}_1, \dots, \mathbf{w}_q\}$ are orthogonal. Moreover, since each vector in $\{\mathbf{w}_1, \dots, \mathbf{w}_q\}$ belongs to $S^\perp$, each of them is orthogonal to every vector in $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$. Thus $\{\mathbf{v}_1, \dots, \mathbf{v}_p, \mathbf{w}_1, \dots, \mathbf{w}_q\}$ is an orthogonal set.

In order to show the spanning property observe that by the Orthogonal Decomposition Theorem any vector $\mathbf{y} \in \mathbb{R}^n$ can be written as a sum $\hat{\mathbf{y}} + \mathbf{z}$ with $\hat{\mathbf{y}} \in S$ and $\mathbf{z} \in S^\perp$. Notice now that there exist scalars $\alpha_1, \dots, \alpha_p$ and $\beta_1, \dots, \beta_q$, such that

$$\begin{aligned} \hat{\mathbf{y}} &= \alpha_1 \mathbf{v}_1 + \dots + \alpha_p \mathbf{v}_p, \\ \mathbf{z} &= \beta_1 \mathbf{w}_1 + \dots + \beta_q \mathbf{w}_q, \end{aligned}$$

because $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ is a spanning set for S and $\{\mathbf{w}_1, \dots, \mathbf{w}_q\}$ is a spanning set for $S^\perp$. Hence

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} = \alpha_1 \mathbf{v}_1 + \dots + \alpha_p \mathbf{v}_p + \beta_1 \mathbf{w}_1 + \dots + \beta_q \mathbf{w}_q,$$

so $\mathbf{y} \in \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_p, \mathbf{w}_1, \dots, \mathbf{w}_q)$. Thus $\{\mathbf{v}_1, \dots, \mathbf{v}_p, \mathbf{w}_1, \dots, \mathbf{w}_q\}$ is a spanning set for $\mathbb{R}^n$.

- (b) By (a), the $p+q$ vectors $\mathbf{v}_1, \dots, \mathbf{v}_p, \mathbf{w}_1, \dots, \mathbf{w}_q$ form a basis for $\mathbb{R}^n$, so $p+q = \dim \mathbb{R}^n = n$. But $\dim S = p$ and $\dim S^\perp = q$, so $\dim S + \dim S^\perp = n$, as claimed.

Exercise* 4. The Gram Schmidt process starts by setting

$$\mathbf{v}_1 = \mathbf{x}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ -1 \end{pmatrix}.$$

The vector $\mathbf{v}_2$ is constructed by subtracting the orthogonal projection of $\mathbf{x}_2$ onto $\text{Span}(\mathbf{v}_1)$ from $\mathbf{x}_2$, that is,

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 = \mathbf{x}_2 - \frac{-6}{3} \mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}.$$

The vector $\mathbf{v}_3$ is constructed by subtracting the orthogonal projection of $\mathbf{x}_3$ onto $\text{Span}(\mathbf{v}_1, \mathbf{v}_2)$ from $\mathbf{x}_3$, that is,

$$\mathbf{v}_3 = \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} \mathbf{v}_2 = \mathbf{x}_3 - \frac{-3}{3} \mathbf{v}_1 - \frac{3}{3} \mathbf{v}_2 = \begin{pmatrix} 0 \\ -1 \\ 1 \\ -1 \end{pmatrix},$$

finally producing the orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ for S .

By the Best Approximation Theorem, the point in S closest to $\mathbf{y}$ is the orthogonal projection $\hat{\mathbf{y}}$ of $\mathbf{y}$ onto S . Using the Orthogonal Decomposition Theorem and the fact that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal basis for S we find

$$\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 + \frac{\mathbf{y} \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} \mathbf{v}_2 + \frac{\mathbf{y} \cdot \mathbf{v}_3}{\mathbf{v}_3 \cdot \mathbf{v}_3} \mathbf{v}_3 = \frac{3}{3} \mathbf{v}_1 + \frac{-3}{3} \mathbf{v}_2 + \frac{3}{3} \mathbf{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -3 \end{pmatrix}.$$

Exercise* 5. In order to write down the normal equations for the least squares problem $A\mathbf{x} = \mathbf{b}$ we first calculate

$$A^T A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix},$$

and

$$A^T \mathbf{b} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \\ 8 \\ 2 \end{pmatrix} = \begin{pmatrix} 14 \\ 4 \\ 10 \end{pmatrix},$$

so the normal equations are

$$\begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 14 \\ 4 \\ 10 \end{pmatrix}.$$

The above system is solved using Gaussian elimination:

$$\left(\begin{array}{ccc|c} 4 & 2 & 2 & 14 \\ 2 & 2 & 0 & 4 \\ 2 & 0 & 2 & 10 \end{array} \right) \sim \left(\begin{array}{ccc|c} 2 & 1 & 1 & 7 \\ 1 & 1 & 0 & 2 \\ 1 & 0 & 1 & 5 \end{array} \right) \sim \left(\begin{array}{ccc|c} 1 & 0 & 1 & 5 \\ 1 & 1 & 0 & 2 \\ 2 & 1 & 1 & 7 \end{array} \right) \sim \left(\begin{array}{ccc|c} 1 & 0 & 1 & 5 \\ 0 & 1 & -1 & -3 \\ 0 & 1 & -1 & -3 \end{array} \right) \sim \left(\begin{array}{ccc|c} 1 & 0 & 1 & 5 \\ 0 & 1 & -1 & -3 \\ 0 & 0 & 0 & 0 \end{array} \right).$$

The leading variables are x_1 and x_2 and the free variable is x_3 . Setting $x_3 = \alpha$, we find $x_2 = -3 + \alpha$ and $x_1 = 5 - \alpha$. Thus the set of least squares solutions of $A\mathbf{x} = \mathbf{b}$ is given by

$$\left\{ \begin{pmatrix} 5 - \alpha \\ -3 + \alpha \\ \alpha \end{pmatrix} \mid \alpha \in \mathbb{R} \right\}.$$

Exercise 6.

- (a) If $A\mathbf{x} = \mathbf{0}$ then $A^T A\mathbf{x} = A^T \mathbf{0} = \mathbf{0}$.
- (b) Suppose that $A^T A\mathbf{x} = \mathbf{0}$. Then $\mathbf{x}^T A^T A\mathbf{x} = \mathbf{x}^T \mathbf{0} = \mathbf{0}$. Thus

$$\mathbf{0} = \mathbf{x}^T A^T A\mathbf{x} = (A\mathbf{x})^T (A\mathbf{x}) = (A\mathbf{x}) \cdot (A\mathbf{x}) = \|A\mathbf{x}\|^2,$$

and hence $A\mathbf{x} = \mathbf{0}$.

- (c) By (a) and (b) above, $A\mathbf{x} = \mathbf{0}$ if and only if $A^T A\mathbf{x} = \mathbf{0}$, so $\mathbf{x} \in N(A)$ if and only if $\mathbf{x} \in N(A^T A)$. Thus $N(A^T A) = N(A)$.
- (d) Note that the columns of A are linearly independent if and only if the only solution of $A\mathbf{x} = \mathbf{0}$ is the zero solution. Thus, the columns of A are linearly independent if and only if $N(A) = \{\mathbf{0}\}$. But, by (c) above $N(A) = N(A^T A)$, so the columns of A are linearly independent if and only if $N(A^T A) = \{\mathbf{0}\}$, which, by the Invertible Matrix Theorem, is the case if and only if $A^T A$ is invertible.

Exercise* 7.

- (a) The characteristic polynomial is given by

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 4 - \lambda & 9 \\ -1 & -2 - \lambda \end{vmatrix} = (4 - \lambda)(-2 - \lambda) + 9 \\ &= -8 - 4\lambda + 2\lambda + \lambda^2 + 9 = \lambda^2 - 2\lambda + 1 = (\lambda - 1)^2, \end{aligned}$$

so the only eigenvalue of the matrix is $\lambda_1 = 1$.

In order to find the corresponding eigenspace we find the nullspace of $A - \lambda_1 I$ using Gaussian elimination:

$$A - \lambda_1 I = \begin{pmatrix} 4 - 1 & 9 \\ -1 & -3 \end{pmatrix} = \begin{pmatrix} 3 & 9 \\ -1 & -3 \end{pmatrix} \sim \begin{pmatrix} 1 & 3 \\ -1 & -3 \end{pmatrix} \sim \begin{pmatrix} 1 & 3 \\ 0 & 0 \end{pmatrix},$$

so setting $x_2 = \alpha$ we find $x_1 = -3x_2 = -3\alpha$. Thus every vector in $N(A - \lambda_1 I)$ is of the form

$$\begin{pmatrix} -3\alpha \\ \alpha \end{pmatrix} = \alpha \begin{pmatrix} -3 \\ 1 \end{pmatrix},$$

so the eigenspace corresponding to the eigenvalue 1 is

$$\left\{ \alpha \begin{pmatrix} -3 \\ 1 \end{pmatrix} \mid \alpha \in \mathbb{R} \right\}.$$

- (b) The characteristic polynomial is given by

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} -4 - \lambda & 0 & 3 \\ -12 & 2 - \lambda & 6 \\ -6 & 0 & 5 - \lambda \end{vmatrix} = (-4 - \lambda) \begin{vmatrix} 2 - \lambda & 6 \\ 0 & 5 - \lambda \end{vmatrix} + 3 \begin{vmatrix} -12 & 2 - \lambda \\ -6 & 0 \end{vmatrix} \\ &= (-4 - \lambda)(2 - \lambda)(5 - \lambda) + 18(2 - \lambda) = (2 - \lambda)[(-4 - \lambda)(5 - \lambda) + 18] \\ &= (2 - \lambda)(\lambda^2 - \lambda - 2) = (2 - \lambda)(\lambda - 2)(\lambda + 1) = -(\lambda - 2)^2(\lambda + 1), \end{aligned}$$

so the eigenvalues of the matrix are $\lambda_1 = -1$ and $\lambda_2 = 2$.

In order to find the eigenspace corresponding to λ_1 we find the nullspace of $A - \lambda_1 I$ using Gaussian elimination:

$$A - \lambda_1 I = \begin{pmatrix} -4 - (-1) & 0 & 3 \\ -12 & 2 - (-1) & 6 \\ -6 & 0 & 5 - (-1) \end{pmatrix} = \begin{pmatrix} -3 & 0 & 3 \\ -12 & 3 & 6 \\ -6 & 0 & 6 \end{pmatrix} \sim \dots \sim \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix},$$

so setting $x_3 = \alpha$ we find $x_2 = 2x_3 = 2\alpha$ and $x_1 = x_3 = \alpha$. Thus, every vector in $N(A - \lambda_1 I)$ is of the form

$$\begin{pmatrix} \alpha \\ 2\alpha \\ \alpha \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix},$$

so the eigenspace corresponding to the eigenvalue -1 is

$$\left\{ \alpha \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \mid \alpha \in \mathbb{R} \right\}.$$

In order to find the eigenspace corresponding to λ_2 we find the nullspace of $A - \lambda_2 I$ using Gaussian elimination:

$$A - \lambda_2 I = \begin{pmatrix} -4-2 & 0 & 3 \\ -12 & 2-2 & 6 \\ -6 & 0 & 5-2 \end{pmatrix} = \begin{pmatrix} -6 & 0 & 3 \\ -12 & 0 & 6 \\ -6 & 0 & 3 \end{pmatrix} \sim \dots \sim \begin{pmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

so setting $x_2 = \alpha$ and $x_3 = \beta$ we find $x_1 = \frac{1}{2}x_3 = \frac{1}{2}\beta$. Thus every vector in $N(A - \lambda_2 I)$ is of the form

$$\begin{pmatrix} \frac{1}{2}\beta \\ \alpha \\ \beta \end{pmatrix} = \alpha \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} \frac{1}{2} \\ 0 \\ 1 \end{pmatrix},$$

and so the eigenspace corresponding to the eigenvalue 2 is

$$\left\{ \alpha \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} \frac{1}{2} \\ 0 \\ 1 \end{pmatrix} \mid \alpha, \beta \in \mathbb{R} \right\}.$$

OB 6/12/2007