MTH5112 Linear Algebra I
MTH5212 Applied Linear Algebra

COURSEWORK 9 — SOLUTIONS

Exercise (∗) 1. The solutions will appear on WeBWork after CW9 due date.

Exercise 2. (a) We have

$$\begin{aligned} ||\mathbf{x}+\mathbf{y}||^2 &= (\mathbf{x}+\mathbf{y})\cdot(\mathbf{x}+\mathbf{y}) \\ &= \mathbf{x}\cdot\mathbf{x}+\mathbf{x}\cdot\mathbf{y}+\mathbf{y}\cdot\mathbf{x}+\mathbf{y}\cdot\mathbf{y} \\ &= ||\mathbf{x}||^2+||\mathbf{y}||^2+2(\mathbf{x}\cdot\mathbf{y}) \end{aligned} \tag{1}$$

Now, by definition, $\mathbf{x}$ and $\mathbf{y}$ are orthogonal if and only if $\mathbf{x}\cdot\mathbf{y}=0$, so it follows from equation (1) that $\mathbf{x}$ and $\mathbf{y}$ are orthogonal if and only if

$$||\mathbf{x}+\mathbf{y}||^2 = ||\mathbf{x}||^2+||\mathbf{y}||^2.$$

(b) First observe that the inequality is clearly true if either of $\mathbf{u}$ or $\mathbf{v}$ are the zero vector, so we can assume that both are *not* the zero vector. In particular, $||\mathbf{u}||\neq 0$ and $||\mathbf{v}||\neq 0$. Taking $\mathbf{x}=||\mathbf{u}||\mathbf{v}$ and $\mathbf{y}=-||\mathbf{v}||\mathbf{u}$ in (1) as per the hint, we obtain

$$\begin{aligned} ||(||\mathbf{u}||\mathbf{v})+(-||\mathbf{v}||\mathbf{u})||^2 &= ||(||\mathbf{u}||\mathbf{v})||^2+||(-||\mathbf{v}||\mathbf{u})||^2+2((||\mathbf{u}||\mathbf{v})\cdot(-||\mathbf{v}||\mathbf{u})) \\ &= ||\mathbf{u}||^2||\mathbf{v}||^2+(-||\mathbf{v}||)^2||\mathbf{u}||^2-2||\mathbf{u}||||\mathbf{v}||\mathbf{v}\cdot\mathbf{u} \\ &= 2||\mathbf{u}||^2||\mathbf{v}||^2-2||\mathbf{u}||||\mathbf{v}||\mathbf{v}\cdot\mathbf{u}. \end{aligned}$$

Because the left-hand side above is non-negative, the right-hand side is also non-negative, so we have

$$0\leq 2||\mathbf{u}||^2||\mathbf{v}||^2-2||\mathbf{u}||||\mathbf{v}||\mathbf{v}\cdot\mathbf{u},$$

or in other words,

$$||\mathbf{u}||||\mathbf{v}||\mathbf{v}\cdot\mathbf{u}\leq ||\mathbf{u}||^2||\mathbf{v}||^2. \tag{2}$$

Because we may assume that $||\mathbf{u}||\neq 0$ and $||\mathbf{v}||\neq 0$ (see above), we can divide both sides of the above inequality by $||\mathbf{u}||||\mathbf{v}||$ to obtain

$$\mathbf{v}\cdot\mathbf{u}\leq ||\mathbf{u}||||\mathbf{v}||.$$

We're not *quite* finished yet, because we want to show that the *absolute value* of $\mathbf{v}\cdot\mathbf{u}$ is less than or equal to $||\mathbf{u}||||\mathbf{v}||$. However, if we run through the above argument again, except with $\mathbf{y}=+||\mathbf{v}||\mathbf{u}$, then we obtain

$$-\mathbf{v}\cdot\mathbf{u}\leq ||\mathbf{u}||||\mathbf{v}||. \tag{3}$$

Putting (2) and (3) together yields the desired inequality, namely

$$|\mathbf{v}\cdot\mathbf{u}|\leq ||\mathbf{u}||||\mathbf{v}||.$$

(c) Equation (1) and the Cauchy–Schwartz inequality (from part (b)) yield

$$\begin{aligned} ||\mathbf{u}+\mathbf{v}||^2 &= ||\mathbf{u}||^2 + ||\mathbf{v}||^2 + 2(\mathbf{u}\cdot\mathbf{v}) \\ &\leq ||\mathbf{u}||^2 + ||\mathbf{v}||^2 + 2|\mathbf{u}\cdot\mathbf{v}| \\ &\leq ||\mathbf{u}||^2 + ||\mathbf{v}||^2 + 2||\mathbf{u}||||\mathbf{v}|| \\ &= (||\mathbf{u}|| + ||\mathbf{v}||)^2. \end{aligned}$$

Now taking square roots of both sides gives the desired inequality, namely

$$||\mathbf{u}+\mathbf{v}|| \leq ||\mathbf{u}|| + ||\mathbf{v}||.$$

Exercise 3. (a) As with all 'subspace' proofs, we must show three things: (i) $H^\perp$ is non-empty, (ii) $H^\perp$ is closed under addition, (iii) $H^\perp$ is closed under scalar multiplication.

(i) Since the zero vector is orthogonal to every vector in $\mathbb{R}^n$, it is, in particular, orthogonal to every vector in H, so the zero vector is an element of $H^\perp$ and hence $H^\perp$ is non-empty.

(ii) Let $\mathbf{x},\mathbf{y}\in H^\perp$. We must show that $\mathbf{x}+\mathbf{y}\in H^\perp$, i.e. we must show that $(\mathbf{x}+\mathbf{y})\cdot\mathbf{v}=0$ for every $\mathbf{v}\in H$. Our strategy is the usual one for these kinds of proofs: use the fact that $\mathbf{x},\mathbf{y}\in H^\perp$, i.e. that $\mathbf{x}\cdot\mathbf{v}=0$ and $\mathbf{y}\cdot\mathbf{v}=0$ for every $\mathbf{v}\in H$. This allows us to say that

$$(\mathbf{x}+\mathbf{y})\cdot\mathbf{v} = \mathbf{x}\cdot\mathbf{v} + \mathbf{y}\cdot\mathbf{v} = 0+0 = 0,$$

which is what we wanted. Hence, $H^\perp$ is indeed closed under addition.

(iii) Let $\mathbf{x}\in H^\perp$ and $\alpha\in\mathbb{R}$. We must show that $\alpha\mathbf{x}\in H^\perp$, i.e. that $(\alpha\mathbf{x})\cdot\mathbf{v}=0$ for every $\mathbf{v}\in H$. Since we are assuming that $\mathbf{x}\cdot\mathbf{v}=0$, we can argue that

$$(\alpha\mathbf{x})\cdot\mathbf{v} = \alpha(\mathbf{x}\cdot\mathbf{v}) = \alpha\cdot 0 = 0,$$

which is what we wanted. Hence, $H^\perp$ is indeed closed under scalar multiplication.

(b) Suppose that $H=\text{span}(\mathbf{v}_1,\ldots,\mathbf{v}_r)$, and let $\mathbf{x}\in\mathbb{R}^n$. If $\mathbf{x}\in H^\perp$ then $\mathbf{x}$ is orthogonal to every vector in H, so in particular $\mathbf{x}$ is orthogonal to the spanning vectors $\mathbf{v}_1,\ldots,\mathbf{v}_r$. It remains to prove the converse, i.e. that if $\mathbf{x}$ is orthogonal to $\mathbf{v}_1,\ldots,\mathbf{v}_r$ then it is orthogonal to *every* vector in H. If $\mathbf{y}$ is any vector in H then we can write

$$\mathbf{y} = \alpha_1\mathbf{v}_1 + \ldots \alpha_r\mathbf{v}_r$$

for some scalars $\alpha_1,\ldots,\alpha_r$. We therefore have

$$\begin{aligned} \mathbf{x}\cdot\mathbf{y} &= \mathbf{x}\cdot(\alpha_1\mathbf{v}_1+\cdots+\alpha_r\mathbf{v}_r) \\ &= \mathbf{x}\cdot(\alpha_1\mathbf{v}_1)+\cdots+\mathbf{x}\cdot(\alpha_r\mathbf{v}_r) \\ &= \alpha_1(\mathbf{x}\cdot\mathbf{v}_1)+\cdots+\alpha_r(\mathbf{x}\cdot\mathbf{v}_r) \\ &= \alpha_1\cdot 0+\cdots+\alpha_r\cdot 0 \\ &= 0+\cdots+0 \\ &= 0. \end{aligned}$$

Hence, $\mathbf{x}$ is indeed orthogonal to the arbitrary vector $\mathbf{y}\in H$, and so $x\in H^\perp$.

(c) Let r denote the dimension of H, and choose a basis $\mathbf{v}_1,\ldots,\mathbf{v}_r$ for H. Form the $r\times n$ matrix A whose rows are the vectors $\mathbf{v}_1^T,\ldots,\mathbf{v}_r^T$. Then, since $\mathbf{v}_1,\ldots,\mathbf{v}_r$ are linearly independent, they form a basis for the row space of A, so we have

$$\text{rank}(A) = r = \text{dim}(H).$$

On the other hand, by part (b), the nullspace of A is equal to $H^\perp$ (i.e. the solutions of $A\mathbf{x}=\mathbf{0}$ are precisely the vectors that are orthogonal to all of $\mathbf{v}_1,\ldots,\mathbf{v}_r$, and by part (b) these are precisely the vectors that are orthogonal to *every* vector in H). Hence, the nullity of A is the dimension of $H^\perp$, i.e.

$$\mathsf{null}(A) = \mathsf{dim}(H^\perp).$$

But now the rank–nullity theorem tells us that

$$n = \mathsf{rank}(A) + \mathsf{null}(A) = \mathsf{dim}(H) + \mathsf{dim}(H^\perp),$$

which is exactly what we wanted to prove.

Exercise 4. (a) Let A be the 2×3 matrix whose rows are $\mathbf{u}^T$ and $\mathbf{v}^T$, i.e.

$$A = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & -3 \end{pmatrix}.$$

Then $H^\perp$ is just the nullspace of A, i.e. the set of solutions of the linear system $A\mathbf{x} = \mathbf{0}$. Since A is already in row echelon form, we can easily find the solution set of this system. Letting α denote the free variable x_3, we find that $x_2 = 3x_3 = 3\alpha$ and $x_1 = x_3 - x_2 = \alpha - 3\alpha = -2\alpha$, so we have

$$H^\perp = N(A) = \{(-2\alpha, 3\alpha, \alpha)^T : \alpha \in \mathbb{R}\} = \mathsf{span}((-2, 3, 1)^T).$$

In other words, $\{(-2, 3, 1)\}$ is a basis for $H^\perp$.

(b) H is a plane through the origin in $\mathbb{R}^3$, and $H^\perp$ is the line through the origin which is perpendicular to this plane.

Exercise 5. (a) The vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly independent — and hence form a basis for the 3-dimensional vector space $\mathbb{R}^3$ — because the matrix with columns $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ has nonzero determinant:

$$\begin{vmatrix} 1 & -4 & 2 \\ 2 & -2 & -2 \\ 2 & 4 & 1 \end{vmatrix} = \begin{vmatrix} 1 & -4 & 2 \\ 0 & 6 & -6 \\ 0 & 12 & -3 \end{vmatrix} \begin{matrix} \\ R_2 \to R_2 - 2R_1 \\ R_3 \to R_3 - 2R_1 \end{matrix} = 1 \cdot \begin{vmatrix} 6 & -6 \\ 12 & -3 \end{vmatrix} = -18 + 72 = 54 \neq 0.$$

To see that $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ form an orthogonal set, we simply check that the three dot products $\mathbf{v}_1 \cdot \mathbf{v}_2$, $\mathbf{v}_1 \cdot \mathbf{v}_3$ and $\mathbf{v}_2 \cdot \mathbf{v}_3$ are all equal to 0:

$$\begin{aligned} \mathbf{v}_1 \cdot \mathbf{v}_2 &= 1 \cdot (-4) + 2 \cdot (-2) + 2 \cdot 4 = -4 - 4 + 8 = 0 \\ \mathbf{v}_1 \cdot \mathbf{v}_3 &= 1 \cdot 2 + 2 \cdot (-2) + 2 \cdot 1 = 2 - 4 + 2 = 0 \\ \mathbf{v}_2 \cdot \mathbf{v}_3 &= (-4) \cdot 2 + (-2) \cdot (-2) + 4 \cdot 1 = -8 + 4 + 4 = 0. \end{aligned}$$

(b) Since $B = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal basis for $\mathbb{R}^3$, Theorem 6.13 from lectures tells us that the coordinate vector of a vector $\mathbf{y} \in \mathbb{R}^3$ with respect to the basis B is

$$[\mathbf{y}]_B = (c_1, c_2, c_3)^T \quad \text{where} \quad c_i = \frac{\mathbf{y} \cdot \mathbf{v}_i}{\mathbf{v}_i \cdot \mathbf{v}_i} \quad \text{for each} \quad i \in \{1, 2, 3\},$$

so we just need to apply this theorem to the given vectors $\mathbf{y} = \mathbf{u}$ and $\mathbf{y} = \mathbf{v}$. We calculate that

$$\begin{aligned} [\mathbf{u}]_B &= \left(\frac{(-1,5,3)^T \cdot (1,2,2)^T}{(1,2,2)^T \cdot (1,2,2)^T}, \frac{(-1,5,3)^T \cdot (-4,-2,4)^T}{(-4,-2,4)^T \cdot (-4,-2,4)^T}, \frac{(-1,5,3)^T \cdot (2,-2,1)^T}{(2,-2,1)^T \cdot (2,-2,1)^T} \right)^T \\ &= \left(\frac{-1+10+6}{1+4+4}, \frac{4-10+12}{16+4+16}, \frac{-2-10+3}{4+4+1} \right)^T \\ &= \left(\frac{5}{3}, \frac{1}{6}, -1 \right)^T \end{aligned}$$

and

$$
\begin{aligned}
[\mathbf{v}]_B &= \left(\frac{(6,-2,2)^T \cdot (1,2,2)^T}{(1,2,2)^T \cdot (1,2,2)^T}, \frac{(6,-2,2)^T \cdot (-4,-2,4)^T}{(-4,-2,4)^T \cdot (-4,-2,4)^T}, \frac{(6,-2,2)^T \cdot (2,-2,1)^T}{(2,-2,1)^T \cdot (2,-2,1)^T} \right)^T \\
&= \left(\frac{6-4+4}{1+4+4}, \frac{-24+4+8}{16+4+16}, \frac{12+4+2}{4+4+1} \right)^T \\
&= \left(\frac{2}{3}, -\frac{1}{3}, 2 \right)^T.
\end{aligned}
$$