SOLUTIONS TO SELECTED ADDITIONAL EXERCISES FOR MATHEMATICS 132 — Part 2

Winter 2004

II. Perpendicularity (Orthogonality)

II.1 : Orthogonal bases

3. An n-**dimensional flag** in $\mathbf{R}^n$ is defined to be a chain of subspaces

$$\mathcal{F} \quad = \quad W_1 \subset W_2 \subset \ \cdots \ \subset W_n = \mathbf{R}^n$$

such that $\dim W_i = i$ for all i. A *compatible orthonormal basis* for $\mathcal{F}$ is an ordered orthonormal basis $\mathcal{U} = \{\, \mathbf{u}_1, \ \cdots \ , \mathbf{u}_n \,\}$ such that each W_i is spanned by the first i vectors in $\mathcal{U}$.

(a) Suppose that we are given a sequence of vectors $\mathbf{a}_i$ such that $\mathbf{a}_1 \neq \mathbf{0}$ and $\mathbf{a}_i \notin W_(i-1)$ for all $i \geq 2$. Prove that the first i vectors of the set $\mathcal{A} = \{\, \mathbf{a}_1, \ \cdots \ , \mathbf{a}_n \,\}$ form a basis for W_i.

SOLUTION.

Let $\mathcal{A}_i$ be the first i vectors in $\mathcal{A}$, If $i = 1$ then $\mathcal{A}_1 = \{\mathbf{a}_1\}$ is a basis for $W(1)$ because $\mathbf{a}_1 \neq \mathbf{0}$ and $W(1)$ is 1-dimensional. Suppose by induction that the result is true for $W(i-1)$, so that $\mathcal{A}_{i-1}$ is a basis for the given subspace. To show that $\mathcal{A}_i$ is a basis for $W(i)$, note that it is enough to show $\mathcal{A}_i$ is a linearly independent because it has the same number of vectors as a basis for $W(i)$. Suppose that $\sum_{j \leq i} \ c_j \, \mathbf{a}_j = \mathbf{0}$. If $c_i = 0$ then by linear independence of $\mathcal{A}_{i-1}$ the remaining coefficients c_j must also be equal to zero. On the other hand, if $c_i \neq 0$ then we may solve for $\mathbf{a}_i$ and write

$$\mathbf{a}_i \quad = \quad \sum_{j<i} \ c_i^{-1} \, c_j \, \mathbf{a}_j$$

and since the latter form a basis for $W(i-1)$ this would imply that $\mathbf{a}_i \in W(i-1)$. But this is false by our hypotheses, and therefore c_i, and all the preceding constants c_j, must be equal to zero; therefore $\mathcal{A}_i$ is a linearly independent.■

(b) Prove that every n-dimensional flag has a compatible orthonormal basis $\mathcal{U}$ and if $\mathcal{V} = \{\, \mathbf{v}_1, \ \cdots \ , \mathbf{v}_n \,\}$ is any other compatible orthonormal basis then for all i we have $\mathbf{v}_i = \varepsilon_i \, \mathbf{u}_i$ where $\varepsilon_i = \pm 1$.

SOLUTION.

The existence follows from the Gram-Schmidt Process, which recursively constructs orthonormal vectors $\mathbf{u}_i$ such that the first i of them span $W(i)$ for each i. Suppose now that we have two compatible orthonormal bases $\mathcal{U}$ and $\mathcal{V}$. We shall show $\mathbf{v}_i = \varepsilon_i \, \mathbf{u}_i$ for all i. If $i = 1$ then this is true because $\mathbf{u}_1, \, \mathbf{v}_1 \in W(1)$ have unit length and each is a scalar multiple of each other; *i.e.*. we have $\mathbf{v}_1 = c \, \mathbf{u}_1$, and therefore $1 = |\mathbf{v}_1| = |c \, \mathbf{u}_1| = |c| \, |\mathbf{u}_1| = |c|$ implies $c = \pm 1$. Suppose now that the result is known for all $\mathbf{v}_k$ such that $k < i$, where $i \geq 2$. Since $\mathbf{v}_i \in W(i)$ we have $\mathbf{v}_i = \sum_{k<i} \ c_k \, \mathbf{u}_k$. The compatibility

hypothesis implies $\langle \mathbf{v}_i, |bfv_k\rangle = 0$ for $i < k$, and since $v_k = \varepsilon_k \mathbf{u}_k$ it follows that $\mathbf{v}_i$ is also orthogonal to $\mathbf{u}_k$. Therefore the formula

$$\mathbf{v}_i \;=\; \sum_{k \le i} \langle \mathbf{v}_i, \mathbf{u}_k \rangle \, \mathbf{u}_k$$

implies that $\mathbf{v}_i$ is a scalar multiple of $\mathbf{u}_i$. One can now repeat the argument in the case $i = 1$ to show that $\mathbf{v}_i = \pm\, \mathbf{u}_i$.■

II.2 : Orthogonal projections and adjoints

1. Let V be a finite dimensional inner product space, and suppose that E_1 and E_2 are self-adjoint idempotents satisfying $E_1 E_2 = E_2 E_1$. Let W_1 and W_2 be the images of E_1 and E_2 respectively.

(a) Prove that $E_1 E_2$ and $E_1 + E_2 - E_1 E_2$ are the perpendicular projections onto $W_1 \cap W_2$ and $W_1 + W_2$ respectively.

SOLUTION.

We first verify the statement regarding $E_1 E_2$. First of all, the latter is a self adjoint idempotent because the commutativity relation $E_1 E_2 = E_2 E_1$ implies that

$$(E_1 E_2)^* = E_2^* E_1^* = E_2 E_1 = E_1 E_2$$

and also that

$$(E_1 E_2)^2 = E_1^2 E_2^2 = E_1 E_2 \; .$$

Next, if $\mathbf{x} \in W_1 \cap W_2$ then $E_1 E_2(\mathbf{x}) = E_1(\, E_2(\mathbf{x})\,) = E_1(\mathbf{x})$ because $\mathbf{x} \in W_2$ and the right hand side is equal to $\mathbf{x}$ because te latter lies in W_1. To complete the proofs of the assertions regarding $E_1 E_2$ we need to prove that the latter sends every vector in $(W_1 \cap W_2)^\perp$ to $\mathbf{0}$. The main step in this argument is to show that $(W_1 \cap W_2)^\perp = W_1^\perp + W_2^\perp$. Since $(W_1 \cap W_2)^\perp$ contains both $W_1^\perp$ and $W_2^\perp$ it follows that it contains their sum. On the other hand, the dimension formulas

$$\dim U \;+\; \dim U^\perp \;=\; \dim V$$

$$\dim U_1 \;+\; \dim U_2 \;=\; \dim(U_1 \cap U_2) \;+\; \dim(U_1 + U_2)$$

combine to imply that

$$\dim(W_1 \cap W_2)^\perp \;=\; \dim W_1^\perp \;+\; \dim W_2^\perp$$

and therefore the subspaces must be equal. Given $\mathbf{y} \in (W_1 \cap W_2)^\perp$, use this identity to write $\mathbf{y} = \mathbf{y}_1 + \mathbf{y}_2$ where $\mathbf{y}_i \in W_i^\perp$. We then have

$$E_1 E_2(\mathbf{y}) \;=\; E_1 E_2\,(\mathbf{y}_1 + \mathbf{y}_2) \;=\; E_1 E_2(\mathbf{y}_1) \;+\; E_1 E_2(\mathbf{y}_2) \;=$$

$$E_2 E_1(\mathbf{y}_1) \;+\; E_1 E_2(\mathbf{y}_2) \;=\; E_2(\mathbf{0}) \;+\; E_1(\mathbf{y}_2) \;=\; \mathbf{0} \;+\; \mathbf{0} \;=\; \mathbf{0}\;.$$

Therefore $E_1 E_2$ is perpendicular projection onto $W_1 \cap W_2$.

Similarly, the verification of the second statement begins by showing that $E_1 + E_2 - E_1 E_2$ is a self adjoint idempotent. The self adjointness follows because

$$(E_1 + E_2 - E_1 E_2)^* \;=\; E_1^* + E_2^* - (E_1 E_2)^* \;=\; E_1^* + E_2^* - E_2^* E_1^* \;=$$

$$E_1 + E_2 - E_2\,E_1 \quad = \quad E_1 + E_2 - E_1\,E_2$$

and the idempotence property follows because $E_1\,E_2 = E_2\,E_1$ implies

$$(E_1 + E_2 - E_1\,E_2)^2 \quad = \quad E_1^2 + E_2^2 + 2\,E_1\,E_2 - 2\,E_1^2\,E_2 - 2E_1\,E_2^2 \quad =$$

$$E_1 + E_2 + (2 + 4 - 1)\,E_1\,E_2 \quad = \quad E_1 + E_2 - E_1\,E_2 \ .$$

To see that this defines orthogonal projection onto $W_1 + W_2$, note that if every vector in W_i may be written as a sum of a vector in $W_1 \cap W_2$ and a vector in W_i that is perpendicular to $W_1 \cap W_2$. Therefore, every vector in $W_1 + W_2$ may be written as a sum $\mathbf{x} = \mathbf{x}_0 + \mathbf{x}_1 + \mathbf{x}_2$, where $\mathbf{x}_0 \in W_1 \cap W_2$, $\mathbf{x}_1 \in W_1$ is orthogonal to $W_1 \cap W_2$ and $\mathbf{x}_2 \in W_2$ is orthogonal to $W_1 \cap W_2$. Therefore if $\mathbf{x} \in W_1 + W_2$ then

$$[E_1 + E_2 - E_1\,E_2]\,(\mathbf{x}) \quad = \quad [E_1 + E_2 - E_1\,E_2]\,(\mathbf{x}_0 + \mathbf{x}_1 + \mathbf{x}_2) \quad =$$

$$E_1(\mathbf{x}_0) + E_1(\mathbf{x}_1) + E_1(\mathbf{x}_2) + E_2(\mathbf{x}_0) + E_2(\mathbf{x}_1) + E_2(\mathbf{x}_2) - E_1\,E_2(\mathbf{x}_0) - E_1\,E_2(\mathbf{x}_1) - E_1\,E_2(\mathbf{x}_2) \quad =$$

$$\mathbf{x}_0) \ + \ \mathbf{x}_1) \ + \ \mathbf{0} \ + \ \mathbf{x}_0 \ + \ \mathbf{0} \ + \ \mathbf{x}_2 \ - \ \mathbf{x}_0 \ - \ \mathbf{0} \ - \ \mathbf{0} \quad = \quad \mathbf{x}_0 \ + \ \mathbf{x}_1 \ + \ \mathbf{x}_2 \quad = \quad \mathbf{x} \ .$$

Furthermore, if $\mathbf{y} \in (W_1 + W_2)^\perp$, then $\mathbf{y} \in W_1^\perp$ and $\mathbf{y} \in W_2^\perp$ so that

$$[E_1 + E_2 - E_1\,E_2]\,(\mathbf{y}) \quad = \quad E_1(\mathbf{y}) \ + \ E_2(\mathbf{y}) \ + \ E_1\,E_2(\mathbf{y}) \quad = \quad \mathbf{0} \ + \ \mathbf{0} \ - \ \mathbf{0} \quad = \quad \mathbf{0}\ .\blacksquare$$

(b) Prove that $E_1\,E_2 = E_1$ if and only if $W_1 \subset W_2$.

SOLUTION.

The given equation and **(a)** imply that orthogonal projection onto $W_1 \cap W_2$ equals orthogonal projection onto W_1, and therefore the image is equal to both $W_1 \cap W_2$ and W_1, so that these subspaces are equal and hence $W_1 \subset W_2$. Conversely $W_1 \subset W_2$ implies $W_1 \cap W_2 = W_1$, so that the orthogonal projections onto these subspaces, which are given by $E_1\,E_2$ and E_1, must also be equal.$\blacksquare$.

2.* If V is a vector space, a linear transformation $T : V \to V$ is *idempotent* if and only if $T^2 = T$, and a linear transformation $S : W \to W$ is an *involution* if and only if $S^2 = I$.

(a) Given two linear transformations S and T from V to itself related by the equation $S = I - 2\,T$, prove that S is an involution if and only if T is idempotent.

SOLUTION.

If $S^2 - I$ and $S = I - 2T$, then $T = \frac{1}{2}(I - S)$, and therefore

$$T^2 \quad = \quad \left[\tfrac{1}{2}\,(I - S)\right]^2 \quad = \quad \tfrac{1}{4}\,\left[I - 2S + S^2\right] \quad = \quad \tfrac{1}{4}\,[2I - 2S] \quad = \quad \tfrac{1}{2}\,[I - 2] \quad = \quad T\ .$$

Conversely, if $T^2 = T$ and $S = I - 2T$, then

$$S^2 \quad = \quad (I - 2T)^2 \quad = \quad I - 4T + 4T^2 \quad = \quad I - 4T + 4T \quad = \quad I\ .\blacksquare$$

(b) Suppose that S is an involution. Prove that every vector in V can be written uniquely in the form

$$\mathbf{v} \quad = \quad \mathbf{v}_+ \ + \ \mathbf{v}_-$$

where $\mathbf{v}_+$ and $\mathbf{v}_-$ are (zero or) eigenvectors for ± 1. [*Hints:* $\mathbf{v} \pm S(\mathbf{v})$ is zero or an eigenvector for ± 1. This is very similar to the argument proving that every continuous function on the real line is the sum of an even function and an odd one.]

SOLUTION.

Let $\mathbf{v} \in V$, and let $\mathbf{v}_{\pm} = \frac{1}{2}(\mathbf{v} \pm S(\mathbf{v}))$. It follows immediately that $\mathbf{v} = \mathbf{v}_+ + \mathbf{v}_-$, where $S(\mathbf{v}_{\pm}) = \pm\,\mathbf{v}_{\pm}$ (write out the details, remembering that $S^2 = I$). Suppose we are given another (possibly different) descomposition of the form $\mathbf{v} = \mathbf{x}_+ + \mathbf{x}_-$, where $S(\mathbf{x}_{\pm}) = \pm\,\mathbf{v}_{\pm}$. Then we have $\mathbf{v}_+ - \mathbf{x}_+ = \mathbf{x}_+ - \mathbf{v}_-$, where the left hand side is zero or an eigenvector for $+1$ and the right hand side is zero or an eigenvector for -1. The only way this can happen is if both sides of the equation are equal to zero, so that $\mathbf{v}_+ = \mathbf{x}_+$ and $\mathbf{x}_+ = \mathbf{v}_-$.■

(c) In the situation of the preceding exercises, prove that the map $T : V \to V$ sending $\mathbf{v}$ to $\mathbf{v}_-$ is an idempotent linear transformation such that $S = I - 2T$.

SOLUTION.

We shal show that if $T_0 = \frac{1}{2}(I - S)$, then $T_0(\mathbf{v}) = \mathbf{v}_-$. The definition for T_0 shows that it is linear. But direct computation shows that

$$T_0(\mathbf{v}) \quad = \quad T_0(\mathbf{v}_+ + \mathbf{v}_-) \quad = \quad \left[\tfrac{1}{2}(I - S)\right](\mathbf{v}_+ + \mathbf{v}_-) \quad = \quad \tfrac{1}{2}(\mathbf{v}_+ + \mathbf{v}_-) \; - \; \tfrac{1}{2}S(\mathbf{v}_+ + \mathbf{v}_-) \quad =$$

$$\tfrac{1}{2}(\mathbf{v}_+ + \mathbf{v}_-) \; - \; \tfrac{1}{2}(\mathbf{v}_+ - \mathbf{v}_-) \quad = \quad \mathbf{v}_- \;.■$$

3. Suppose that V is a finite dimensional inner product space and $T : V \to V$ is a linear transformation; as usual, let T^* denote the adjoint of T.

(a) Prove that T is invertible if and only if T^* is invertible.

SOLUTION.

Suppose that T is invertible with inverse S. Then $S\,T = I = T\,S$, and therefore we have

$$T^*\,S^* \quad = \quad (S\,T)^* \quad = \quad I^* = (T\,S)^* \quad = \quad S^*\,T^*$$

and since $I^* = I$ it follows that S^* is inverse to T^*. Similarly if T^* is invertible and its inverse is S^0, then S_0^* is inverse to $(T^*)^* = T$.■

(b) Explain why the ranks of T and T^* are equal.

SOLUTION.

If $\mathcal{U}$ is an ordered orthonormal basis for V and A is the matrix of T with respect to $\mathcal{U}$, then the matrix of T^* with respect to $\mathcal{U}$ is equal to ${}^{\mathbf{T}}\!A$. The ranks of T and T^* are then equal to the column ranks of A and its transpose, which are respectively the column and row ranks of A. Since row and column ranks are equal, this implies that the ranks of T and T^* must be equal.■

(b) Prove that the rank of T is is equal to the ranks of both $T\,T^*$ and $T^*\,T$ (hence all three ranks are equal).

SOLUTION.

If $S : V \to V$ is a linear transformation then $\mathrm{rank}(S) + \mathrm{nullity}(S) = \dim V$, (compare $4(b)$ below), so it is enough to show that the nullities of all three linear transformations are equal; *i.e.*, their kernels have the same dimensions. In the case of T and T^*T, we shall show that the kernels are equal. If $\mathbf{x}$ lies in the kernel of T then $T^*T(\mathbf{x}) = T^*(\,T(\mathbf{x})\,) = T^*(\mathbf{0}) = \mathbf{0}$ shows that $\mathrm{Kernel}(T)$ is contained in $\mathrm{Kernel}(T^*)$. Conversely, if $\mathbf{x}$ lies in the kernel of T^*T then $T^*T(\mathbf{x}) = \mathbf{0}$ implies

$$0 \quad = \quad \langle T^*T(\mathbf{x}),\, \mathbf{x}\rangle \quad = \quad \langle T(\mathbf{x}),\, T(\mathbf{x})\rangle \quad = \quad |T(\mathbf{x})|^2$$

implies that $T(\mathbf{x}) = \mathbf{0}$ and hence that $\mathbf{x}$ lies in the kernel of T. Since they have the same kernels, the linear transformations T and T^*T also have the same ranks. If we apply the same argument to T^* we see that T^* and $T\,T^*$ also have the same kernels and ranks. Since T and T^* also have the same rank, it follows that the ranks of T, T^*, $T\,T^*$ and $T^*\,T$ are all equal.■

4. Let V and W be finite dimensional inner product spaces, let $T : V \to W$ be a linear transformation, and let T^* be the adjoint of T.

(a) Prove that the kernel of T^* is the orthogonal complement of the image of T and the image of T^* is the orthogonal complement of the kernel of T.

SOLUTION.

Suppose that $\mathbf{w}$ lies in the kernel of T^*. Then $T^*(\mathbf{w}) = \mathbf{0}$ and hence

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

for all $\mathbf{v} \in W$. Therefore $\mathbf{w}$ lies in the orthogonal complement of the image of T. Conversely, if $\mathbf{w}$ lies in this orthogonal complement then for all $\mathbf{v} \in V$ we have

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

and since the latter is valid for all $\mathbf{v}$ it follows that $T^*(\mathbf{w}) = \mathbf{0}$; *i.e*, $\mathbf{w}$ lies in the kernel of T^*.

If we apply this to T^* we see that the kernel of $(T^*)^* = T$ is the orthogonal complement of the image of T^*, and hence the orthogonal complements of these respective subspaces are also equal. But the orthogonal complements are merely the orthogonal complement of the kernel of T and the (orthogonal complement of the orthogonal complement of) the image of T^*.■

(b) In analogy with page 138 of the text we shall define the *nullity* of T to be the dimension of the kernel of T, and as in the matrix case we have the formula $\mathrm{rank}(T) + \mathrm{nullity}(T) = \dim V$. Prove that

$$\dim V - \dim W = \mathrm{nullity}(T) - \mathrm{nullity}(T^*) \ .$$

SOLUTION.

We know that $\mathrm{rank}(T) + \mathrm{nullity}(T) = \dim V$, and if we apply this result to T^* we also see that $\mathrm{rank}(T^*) + \mathrm{nullity}(T^*) = \dim W$. Subtracting the first equation from the second, we obtain

$$\dim V - \dim W \;\;=\;\; \mathrm{rank}(T) \;+\; \mathrm{nullity}(T) \;-\; \mathrm{rank}(T^*) \;-\; \mathrm{nullity}(T^*) \;\;=\;\; \mathrm{nullity}(T) - \mathrm{nullity}(T^*)$$

where the final equation holds because the ranks of T and T^* are equal.■

II.3 : Orthogonal matrices

1. Let A be an orthogonal 2×2 matrix.

(a) Using Exercise 39 on page 359 of the text, show that if $\det A = -1$ then A has $\pm, 1$ as eigenvalues and an orthonormal basis of eigenvectors. [*Hint:* First show that the characteristic polynomial is $t^2 - 1$. Why are eigenvectors for $+1$ and -1 perpendicular to each other?]

SOLUTION.

According to the exercise cited in the problem, an orthogonal 2×2 matrix A with determinant -1 has the form

$$\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$$

for some real number θ. Direct calculation shows that the characteristic polynomial of such a matrix is equal to $t^2 - 1$. Since this polynomial has the two distinct roots ± 1 it follows that A is diagonalizable, with a basis of eigenvectors given by $\mathbf{v}_+$ and $\mathbf{v}_-$ (associated to the eigenvalues $+1$ and -1 respectively). We claim that these eigenvectors are orthogonal. Observe that A is symmetric, so that the associated linear transformation $\mathcal{L}_A$ is self adjoint. Then we have

$$\langle \mathbf{v}_+, \mathbf{v}_- \rangle \quad = \quad \langle A(\mathbf{v}_+), \mathbf{v}_- \rangle \quad = \quad \langle \mathbf{v}_+, A(\mathbf{v}_-) \rangle \quad = \quad \langle \mathbf{v}_+, = \mathbf{v}_- \rangle \quad = \quad - \langle \mathbf{v}_+, \mathbf{v}_- \rangle$$

and since 0 is the only real number equal to its own negative these equations imply that $\langle \mathbf{v}_+, \mathbf{v}_- \rangle = 0$. This shows that we have an orthogonal basis of eigenvectors, and we can convert it to an orthonormal basis by taking $\mathbf{v}_+ = |\mathbf{v}_+|^{-1} \cdot \mathbf{v}_+$ and $\mathbf{v}_- = |\mathbf{v}_-|^{-1} \cdot \mathbf{v}_-$.$\blacksquare$

2.** Prove the Recognition Principle for orthogonal linear transformations that was stated in the notes: *Let V be a finite-dimensional inner product space, and let $T : V \to V$ be an zero-preserving isometric map — in other words, $T(\mathbf{0}) = \mathbf{0}$, every point of V has the form $T(\mathbf{x})$ for some $\mathbf{x}$, and $|\mathbf{x} - \mathbf{y}| = |T(\mathbf{x}) - T(\mathbf{y})|$ for all $\mathbf{x}, \mathbf{y} \in V$. Then T is an orthogonal linear transformation.* Here are the basic steps in the recommended approach:

(a) Show that T is 1–1. [*Hint:* Why is the distance between $T(\mathbf{x})$ and $T(\mathbf{y})$ positive if $\mathbf{x} \neq \mathbf{y}$?]

SOLUTION.

If $T(\mathbf{x}) = T(\mathbf{y})$, then $\mathbf{0} = T(\mathbf{x}) = T(\mathbf{y})$ implies

$$|\mathbf{x} - \mathbf{y}| \quad = \quad |T(\mathbf{x}) - T(\mathbf{y})| \quad = \quad \mathbf{0}$$

so that $\mathbf{x} = \mathbf{y}$.$\blacksquare$

(b) Show that $|\mathbf{x}| = |T(\mathbf{x})|$ for all $\mathbf{x} \in V$ [*Hint:* We have $|\mathbf{x} - \mathbf{y}| = |T(\mathbf{x}) - T(\mathbf{y})|$ and $T(\mathbf{0}) = \mathbf{0}$).]

SOLUTION.

As indicated in the hint, we have $|\mathbf{x}| = |\mathbf{x} - \mathbf{0}| = |T(\mathbf{x}) - T(\mathbf{0})| = |T(\mathbf{x})|$.$\blacksquare$

(c) Show that T preserves inner products, and in particular that T sends orthonormal sets into orthonormal sets, and likewise for orthonormal bases.

SOLUTION.

For all vectors $\mathbf{u}$ and $\mathbf{v}$ in V we have the identity

$$\langle \mathbf{u}, \mathbf{v} \rangle \quad = \quad \frac{1}{2}\Big(|\mathbf{u} + \mathbf{v}|^2 - |\mathbf{u}|^2 - |\mathbf{v}|^2 \Big)$$

and if we apply this to the ordered pairs of vectors $(\mathbf{x}, \mathbf{y})$ and $\big(T(\mathbf{x}), T(\mathbf{y})\big)$ we obtain the following equations:

$$\langle \mathbf{x}, \mathbf{y} \rangle \quad = \quad -\frac{1}{2}\Big(|\mathbf{x} - \mathbf{y}|^2 - |\mathbf{x}|^2 - |\mathbf{y}|^2 \Big)$$

$$\langle T(\mathbf{x}),\, T(\mathbf{y})\rangle \quad = \quad -\frac{1}{2}\left(|T(\mathbf{x}) - T(\mathbf{y})|^2 - |T(\mathbf{x})|^2 - |T(\mathbf{y})|^2\right)$$

By the preceding step we know that the expressions on the right hand sides of these equations are equal to each other, and thus the same is true for the left hand sides. This proves that T preserves inner products, and in particular T sends orthogonal pairs of vectors into pairs of the same type. Since T also preserves lengths, it follows that it sens orthonormal sets into orthonormal sets, and since it is $1-1$ it also sends orthonormal sets with n vectors ($i.e.$, bases) into similar sets.■

(d) Let $\mathcal{U}$ be an orthonormal basis for V, and let $\mathcal{U}'$ be its image under T. Express $\mathbf{x}$ and $T(\mathbf{x})$ as explicit linear combinations of the vectors in $\mathcal{U}$ and $\mathcal{U}'$ respectively using the inner product, and use the preceding step to show that T must be linear.

SOLUTION.

Let $\{\mathbf{u}_1,\ \cdots\ ,\mathbf{u}_n\}$ be the orthonormal basis; then we have

$$\mathbf{x} \quad = \quad \sum_{i=1}^{n} \langle \mathbf{x}, \mathbf{u}_i\rangle \cdot \mathbf{u}_i$$

and likewise we have

$$T(\mathbf{x}) \quad = \quad \sum_{i=1}^{n} \langle T(\mathbf{x}), T(\mathbf{u}_i)\rangle \cdot T(\mathbf{u}_i) \ .$$

Since T preserves inner products we know that

$$\langle \mathbf{x}, \mathbf{u}_i\rangle \quad = \quad \langle T(\mathbf{x}), T(\mathbf{u}_i)\rangle$$

for all i, and therefore we have

$$T(\mathbf{x}) \quad = \quad \sum_{i=1}^{n} \langle \mathbf{x}, \mathbf{u}_i\rangle \cdot T(\mathbf{u}_i) \ .$$

Since the inner product is linear it follows that each summand on the right hand side of the preceding equation defines a linear transformation from V to itself, and therefore the entire sum also defines a linear transformation from V to itself.■

III. Change of bases

III.A : Review topics

Additional exercises

1. For all scalars a and b the 3×3 matrix

$$A \quad = \quad \begin{pmatrix} 2 & a & 0 \\ 0 & -1 & b \\ 0 & 0 & +1 \end{pmatrix}$$

has a basis of eigenvectors. Find an invertible matrix P such that $P^{-1}AP$ is diagonal.

SOLUTION.

Direct calculation shows that the characteristic polynomial of A is equal to $(2-t)(t^2-1)$, and therefore A has a basis of eigenvectors associated to the three eigenvalues 2, -1 and $+1$. Therefore we need to find the null spaces of the following three matrices:

$$A - 2I = \begin{pmatrix} 0 & a & 0 \\ 0 & -3 & b \\ 0 & 0 & -1 \end{pmatrix} \quad , \quad A + I = \begin{pmatrix} 3 & a & 0 \\ 0 & 0 & b \\ 0 & 0 & 2 \end{pmatrix} \quad , \quad A - I = \begin{pmatrix} 1 & a & 0 \\ 0 & -2 & b \\ 0 & 0 & 0 \end{pmatrix}$$

In the first case the null space is spanned by $\mathbf{e}_1$, in the second case the null space is spanned by $a\,\mathbf{e}_1 - 3\,\mathbf{e}_2$, and in the third case the null space is spanned by $-a\,b\,\mathbf{e}_1 + b\,\mathbf{e}_2 + 2\,\mathbf{e}_3$. The invertible matrix P may be taken to be any matrix whose columns are nonzero multiples of these three vectors. In particular, if we choose P such that the three vectors listed above are its columns in the given order, then it follows that the resulting diagonal matrix has entries 2, -1 and $+1$ in that order.■

III.1 : Similarity of matrices

1. Let A and B be $n \times n$ matrices.

(a) If B is invertible, show that AB is similar to BA.

SOLUTION.

This is a consequence of the following chain of equations:

$$BA = BA(BB^{-1}) = B(AB)B^{-1} \quad ■$$

(b) Give examples of noninvertible $n \times n$ matrices such that AB and BA are not similar. [*Hint:* Why is it enough to find 2×2 matrices A and B such that $AB = 0$ but $BA \neq 0$?]

SOLUTION.

If we can find matrices as in the hint, then AB cannot be similar to BA because the only matrix which is similar to the zero matrix is the zero matrix itself (*Proof:* $P^{-1}\,0\,P = P^{-1}\,0 = 0$). Perhaps the easiest specific examples are given by the 2×2 matrices that have exactly one entry equal to 1 and all other entries equal to 0. In particular, the matrices

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad , \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

satisfy $AB = 0$ but $BA \neq 0$.■

2. Let A be an orthogonal 3×3 matrix.

(a) Prove that there is an orthogonal matrix P such that $P^{-1}AP$ has the form

$$\begin{pmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & \varepsilon \end{pmatrix}$$

where the 2×2 matrix in the upper left hand corner is orthogonal and $\varepsilon = \pm 1$. [*Hint:* Find an ordered orthonormal basis such that the third vector $\mathbf{v}_3$ is an eigenvector for A and $\mathcal{L}_A$ takes the 2-dimensional subspace W of vectors perpendicular to $\mathbf{v}_3$ into itself.]

SOLUTION.

The matrix has an eigenvector, which we may as well assume has unit length. Let $\mathbf{u}_3$ denote this eigenvector. What is the associated eigenvalue? Since A is orthogonal we know $A(\mathbf{u}_3)$ is also a unit vector, and therefore $A(\mathbf{u}_3) = c\,\mathbf{u}_3$ implies that $c = \pm 1$. Let W be the subspace of all vectors perpendicular to $\mathbf{u}_3$. We claim that if $\mathbf{w} \in W$ then $A(\mathbf{w})$ also lies in W. Now $\mathbf{w} \in W$ and $A(\mathbf{u}_3) = \pm\,\mathbf{u}_3$ imply

$$\langle A(\mathbf{w}),\, \mathbf{u}_3\rangle \quad = \quad \langle A(\mathbf{w}),\, \pm\, A(\mathbf{u}_3)\rangle \quad = \quad \langle \mathbf{w},\, \pm\,\mathbf{u}_3\rangle \quad = \quad \pm\,\langle \mathbf{w},\, \mathbf{u}_3\rangle \quad = \quad 0$$

and consequently $A(\mathbf{w}) \in W$ as desired. If we take an orthonormal basis $\{\,\mathbf{u}_1,\, \mathbf{u}_2\,\}$ for W, then the matrix of A with respect to the orthonormal basis $\{\,\mathbf{u}_1,\, \mathbf{u}_2\, \mathbf{u}_3\,\}$ will have the form described in the exercise.■

(b) If $\det A = 1$ does it follow that one can find a similar matrix of the above form with $\varepsilon = +1$? Either prove this or give a counterexample.

SOLUTION.

The answer is YES for the following reason. The lower right entry of the matrix is ± 1, so we only need to consider the case where the entry is -1 and find a new matrix similar to the original ones such for which the lower right entry is $+1$. Suppose that we have a matrix of the form described but the lower right hand entry is -1. Since the determinant of the entire matrix is $+1$ this means that the determinant of the matrix in the upper left hand corner is also equal to -1. By an additional exercise for Section II.3, the matrix in the upper left hand corner has an orthonormal basis of eigenvectors. The eigenvalues for this matrix are ± 1. It follows that the original matrix A is diagonalizable with a 2-dimensional subspace of eigenvectors for the eigenvalue -1 and a 1-dimensional subspace of eigenvectors for the eigenvalue $+1$. If we order this basis properly, it follows that the matrix of A with respect to this ordered basis is diagonal, and in order the diagonal entries are -1, -1 and $+1$.■

3. Suppose that we are given an invertible $n \times n$ matrix A and we write it as $Q\,R$ where Q is orthogonal and R is lower triangular with positive entries down the diagonal. Why is $R\,Q$ similar to $Q\,R$? [*Note:* This simple observation is the basis for a very simple and effective algorithm that can be used to compute the eigenvalues of certain matrices. One can write $R\,Q = Q_1\,R_1$ by the **QR** decomposition theorem from the preceding unit, then one can write the similar matrix $R_1\,Q_1$ as $Q_2\,R_2$, and in favorable cases the resulting sequence of mutually similar matrices converges to a diagonal matrix. In cases that are not so favorable, one can sometimes modify the method slightly in order to obtain the desired eigenvalues.]

SOLUTION.

This follows from the previous observation that $A\,B$ and $B\,A$ are always similar if A and B are invertible.■

III.2 : Invariants of similarity

1. Given an $n \times n$ matrix A with entries $a_{i,j}$, define its *trace*, denoted by $\mathrm{tr}(A)$, to be the scalar $\sum_i\, a_{i,i}$.

(a) Verify that the trace defines a linear transformation from the space of all $n \times n$ matrices to the real numbers (in the terminology of some books, such a map is also called a linear functional).

SOLUTION.

We have $\mathrm{tr}(c\,A) = \sum_i \; c\,a_{i,i} = c\;(\sum_i \; a_{i,i}) = c\,\mathrm{tr}(A)$, and also

$$\mathrm{tr}(A+B) = \sum_i \;(a_{i,i} + b_{i,i}) \quad = \quad \sum_i \; a_{i,i} \;+\; \sum_i \; b_{i,i} \quad = \quad \mathrm{tr}(A) \;+\; \mathrm{tr}(B)$$

and therefore the trace is a linear mapping.■

(b) Prove that $\mathrm{tr}(A\,B) = \mathrm{tr}(B\,A)$ for all $n \times n$, matrices A and B, and prove that similar matrices have the same trace.

SOLUTION.

Write out the traces for the products $A\,B$ and $B\,A$ explicitly:

$$\mathrm{tr}(A\,B) \quad = \quad \sum_i \; \left(\sum_j \; a_{i,j}\,b_{j,i}\right)$$

$$\mathrm{tr}(B\,A) \quad = \quad \sum_i \; \left(\sum_j \; b_{i,j}\,a_{j,i}\right)$$

To see that the expressions on the right hand sides of these equations are equal, make the change of summation variables $k = i$ and $\ell = j$ in the first summation and $k = j$ and $\ell = i$ in the second one.

One can now use the equation $\mathrm{tr}(A\,B) = \mathrm{tr}(B\,A)$ to prove similar matrices have the same trace exactly the same way as one can use the equation $\det(A\,B) = \det(B\,A)$ to prove similar matrices have the same determinant:

$$\mathrm{tr}(P^{1}\,A\,P) \quad = \quad \mathrm{tr}(A\,P\,P^{-1}) \quad = \quad \mathrm{tr}(A\,I) \quad = \quad \mathrm{tr}(A) \;\;■$$

(c) If A is diagonalizable, explain why its trace is equal to $\sum_\lambda \;\lambda\,n(\lambda)$, where λ runs through all the eigenvalues of A and $n(\lambda)$ is the dimension of the eigenspace for λ.

SOLUTION.

Let D be a diagonal matrix such that A is similar to D; then there is a basis of vectors $\mathbf{v}_i$ such that $A(\mathbf{v}_i) = d_{i,i}\,\mathbf{v}_i$. By the previous portions of this exercise the trace of A is equal to the sum of the diagonal entries $d_{i,i}$. For each number λ that appears as a diagonal entry, let $n(\lambda)$ denote the number of times that it appears. It then follows that the $n(\lambda)$ is the dimension of the subspace of eigenvectors associated to λ and that the traces of A and D are equal to $\sum_\lambda \;\lambda\,n(\lambda)$.■

2. Explain why the matrix function $\varphi(A,B) = \mathrm{tr}\left({}^{\mathbf{T}}B\,A\right)$ defines an inner product on the space of $n \times n$ matrices.

SOLUTION.

The formula for the trace of the product from the preceding exercise shows it is equal to

$$\sum_{i,j} a_{i,j}\,b_{i,j} \;.■$$

IV. Complex linear algebra

IV.1 : Complex numbers

1. Using the formulas $e^{ix} = \cos x + i\,\sin x$ and $e^{i5y} = (e^{iy})^5$, derive the following trigonometric identities:

$$\cos 5x \;=\; \cos^5 x \;-\; 10\,\cos^3 x\cdot\sin^2 x \;+\; 5\,\cos x\cdot\sin^4 x$$

$$\sin 5x \;=\; \sin^5 x \;-\; 5\,\sin^3 x\cdot\cos^2 x \;+\; 10\,\sin x\cdot\cos^4 x$$

SOLUTION.

The Binomial Formula implies that

$$(p+q)^5 \;=\; p^5 \;+\; 5\,p^4\,q \;+\; 10\,p^3\,q^2 \;+\; 10\,p^2\,q^3 \;+\; 5\,p\,q^4 \;+\; q^5$$

and if we substitute $\cos x + i\,\sin x$ for $p+q$ we obtain the following:

$$\cos 5x \;+\; i\,\sin 5x \;=\; (\cos\,x \;+\; \sin x)^5 \;=$$

$$\cos^5 x \;+\; 5\,i\,\cos^4 x\,\sin x \;+\; 10\,i^2\,\cos^3 x\,\sin^2 x \;+\; 10\,i^3\,\cos^2 x\,\sin^3 x \;+\; 5\,i^4\,\cos x\,\sin^4 x \;+\; i^5\,\sin^5 x$$

Using the fact that $i^2 = -1$ we may rewrite the right hand side as follows:

$$\left(\cos^5 x \;-\; 10\,\cos^3 x\cdot\sin^2 x \;+\; 5\,\cos x\cdot\sin^4 x\right) \;+\; i\,\left(\sin^5 x \;-\; 5\,\sin^3 x\cdot\cos^2 x \;+\; 10\,\sin x\cdot\cos^4 x\right)$$

This is equal to the left hand side of the original equation, which was $\cos 5x + i\,\sin 5x$, and the formulas in the exercise follow by taking the real and imaginary parts of these expressions.■