Solutions to the selected problems (Homework 5–7)

Linear Algebra

Fall 2010

Page 86, 3) Let \( T \) be the given function, so \( T(x, y, z, t) = \begin{pmatrix} t + x & y + iz \\ y - iz & t - x \end{pmatrix} \).

Then

\[
T(c(x, y, z, w) + (x', y', z', w')) = T(cx + x', cy + y', cz + z', cw + w')
= \begin{pmatrix} ct + t' + cx + x' & cy + y' + i(cz + z') \\ cy + y' - i(cz + z') & ct + t' - (cx + x') \end{pmatrix}
= c \begin{pmatrix} t + x & y + iz \\ y - iz & t - x \end{pmatrix} + \begin{pmatrix} t' + x' & y' + iz' \\ y' - iz' & t' - x' \end{pmatrix}
= cT(x, y, z, w) + T(x', y', z', w'),
\]

for any \( c \in \mathbb{R} \), so \( T \) is linear. To show that \( T \) is an isomorphism, it is enough to show that \( T \) is one-one and onto.

If \( T(x, y, z, w) = 0 \), then \( t + x = y + iz = y - iz = t - x = 0 \), so \( t = x = z = w = 0 \), so \( T \) is one-one. If \( A \) is a Hermitian matrix, then

\[
A = \begin{pmatrix} a & b + ic \\ b - ic & d \end{pmatrix}
\]

where \( a, b, c, d \in \mathbb{R} \). If we let \( t = \frac{a + d}{2}, x = \frac{a - d}{2}, y = b, z = c \), then \( T(x, y, z, w) = A \), so \( T \) is onto.

Page 96, 12. (b) We have

\[
T^m(\alpha_j) = \begin{cases} 
\alpha_{j+m} & \text{if } j \leq n - m \\
0 & \text{if } j > n - m
\end{cases}
\]

So \( T^n(\alpha_i) = 0 \) for every \( 1 \leq i \leq n \), and since every vector can be written as a linear combination of the \( \alpha_i \), \( T^n(\alpha) = 0 \) for every vector \( \alpha \in V \). We have \( T^{n-1}(\alpha_1) = \alpha_n \neq 0 \), so \( T^{n-1} \neq 0 \).
(c) Since $S^{n-1} \neq 0$, we can choose a vector $\alpha$ such that $S^{n-1}(\alpha) \neq 0$. Let $\alpha_1 = \alpha$, $\alpha_2 = S(\alpha), \ldots, \alpha_i = S^{i-1}(\alpha), \ldots, \alpha_n = S^{n-1}(\alpha)$. Clearly $S(\alpha_j) = \alpha_{j+1}$ if $j < n$, and $S(\alpha_n) = 0$ since $S^n = 0$. We claim the $\alpha_j$ are linearly independent. Note that $\alpha_n = S^{n-1}(\alpha) \neq 0$. Assume on the contrary that there is a non-trivial linear relation

$$c_1\alpha_1 + \cdots + c_n\alpha_n = 0$$

(so there is at least one $c_i$ which is not equal to zero). Assume that $t$ is the smallest integer such that $c_t \neq 0$. So we have

$$c_t\alpha_t + \cdots + c_n\alpha_n = 0, \quad c_t \neq 0.$$ 

Then

$$S^{n-t}(c_t\alpha_t + \cdots + c_n\alpha_n) = 0.$$ 

So

$$c_tS^{n-t}(\alpha_t) + \cdots + c_nS^{n-t}(\alpha_n) = 0,$$

but $S^{n-t}(\alpha_{t+1}) = \cdots = S^{n-t}(\alpha_n) = 0$ and $S^{n-t}(\alpha_t) = \alpha_t$ by definition, so

$$c_t\alpha_t = 0.$$ 

But $\alpha_t$ is non-zero by our assumption, so $c_t = 0$, a contradiction. Therefore, there is no non-trivial linear relation between the $\alpha_i$. Thus they are linearly independent and hence form a basis.

(d) Assume that $M^{n-1} \neq 0$ and $M^n = 0$. Define a linear transformation

$$S : F^{n \times 1} \rightarrow F^{n \times 1}$$

such that $S(X) = MX$. Then $S^n(X) = M^nX = 0$, so $S^n = 0$, and $S^{n-1} \neq 0$. So by part (b) there is a basis $B = \{\alpha_1, \ldots, \alpha_n\}$ for $F^{n \times 1}$ such that $S(\alpha_i) = \alpha_{i+1}$ for $1 \leq i \leq n-1$ and $S(\alpha_n) = 0$. The matrix of $S$ in this basis is

$$[S]_B = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

Since $M$ is the matrix of $S$ with respect to the standard basis, $M$ is similar to the above matrix. The same argument shows that $N$ is similar to the above matrix. Since being similar is an equivalent relations, $M$ and $N$ are similar.
Page 106, 11) If $W_1$ and $W_2$ are two subspaces of a vector space $V$, then clearly

$$W_1 \subset W_2 \text{ implies } (W_2)^0 \subset (W_1)^0.$$ 

(a) Since $W_1, W_2 \subset W_1 + W_2$, $(W_1 + W_2)^0 \subset W_1^0$ and $(W_1 + W_2)^0 \subset W_2^0$, so $(W_1 + W_2)^0 \subset W_1^0 \cap W_2^0$. Conversely, assume $f \in W_1^0 \cap W_2^0$, and let $\alpha \in W_1 + W_2$, then $\alpha = \alpha_1 + \alpha_2$ for some $\alpha_1 \in W_1$ and $\alpha_2 \in W_2$. So $f(\alpha) = f(\alpha_1 + \alpha_2) = f(\alpha_1) + f(\alpha_2) = 0$. Therefore, $f \in (W_1 + W_2)^0$.

(b) Since $W_1 \cap W_2 \subset W_1, W_2$, we have

$$W_1^0, W_2^0 \subset (W_1 \cap W_2)^0.$$ 

Since by definition, $W_1^0 + W_2^0$ is the intersection of all subspaces which contain both $W_1^0$ and $W_2^0$, the above inclusion implies that $W_1^0 + W_2^0 \subset (W_1 \cap W_2)^0$. To show that

$$(W_1 \cap W_2)^0 = W_1^0 + W_2^0,$$

it is enough to show that $\dim(W_1 \cap W_2)^0 = \dim(W_1^0 + W_2^0)$ (because a proper subspace of a vector space has dimension smaller than the dimension of the vector space). We have

$$\dim(W_1^0 + W_2^0) = \dim W_1^0 + \dim W_2^0 - \dim(W_1^0 \cap W_2^0) \quad \text{(by Thm. 6, page 46)}$$

$$= \dim W_1^0 + \dim W_2^0 - \dim(W_1 + W_2)^0 \quad \text{(by part (a))}$$

$$= (\dim V - \dim W_1) + (\dim V - \dim W_2) \quad \text{(by Thm. 16, page 101)}$$

$$- (\dim V - \dim(W_1 + W_2))$$

$$= \dim V - (\dim W_1 + \dim W_2 - \dim(W_1 + W_2))$$

$$= \dim V - \dim(W_1 \cap W_2)$$

$$= \dim(W_1 \cap W_2)^0$$

Page 106, 12) Assume $\dim W = r$ and $\dim V = n$. Pick a basis $\alpha_1, \ldots, \alpha_r$ for $W$, and extend it to a basis: $\alpha_1, \ldots, \alpha_n$ for $V$. We know that a linear functional $g : V \to F$ is uniquely determined by its values at the $\alpha_i$. And we also know that for any choice of scalars $a_1, \ldots, a_n \in F$, there is a linear functional $V \to F$ which sends $\alpha_i$ to $a_i$ (such a linear functional is given by $g(c_1 \alpha_1 + \cdots + c_n \alpha_n) = c_1 a_1 + \cdots + c_n a_n$).

Now given $f : W \to F$, define $g$ as follows: $g(\alpha_1) = f(\alpha_1), \ldots, g(\alpha_r) = f(\alpha_r), g(\alpha_r + 1) = 0, \ldots, g(\alpha_n) = 0$. Then we can extend $g$ to the whole $V$. 3
Any vector $\alpha \in V$ can be written uniquely as

$$\alpha = c_1\alpha_1 + \ldots + c_n\alpha_n,$$

and $g(\alpha) = c_1 f(\alpha_1) + \cdots + c_r f(\alpha_r)$. Then $g$ is of course a linear functional.

And if $\alpha$ is already in $W$, then when we write $\alpha$ as above, we have

$$\alpha = c_1\alpha_1 + \ldots + c_r\alpha_r,$$

so

$$g(\alpha) = c_1 f(\alpha_1) + \cdots + c_r f(\alpha_r) = f(c_1\alpha_1 + \cdots + c_r\alpha_r) = f(\alpha).$$

So on $W$, $f = g$.

Page 106 13) We have $h(\alpha) = f(\alpha)g(\alpha)$, so for every $c \in F$,

$$ch(\alpha) = cf(\alpha)g(\alpha).$$

On the other hand,

$$ch(\alpha) = h(c\alpha) = f(c\alpha)g(c\alpha) = cf(\alpha)cg(\alpha) = c^2f(\alpha)g(\alpha).$$

Comparing the above two equalities, we see for every $c \in F$, and $\alpha \in V$:

$$cf(\alpha)g(\alpha) = c^2f(\alpha)g(\alpha).$$

Pick an arbitrary $c \neq 0, 1$. For every $\alpha \in V$, we have

$$f(\alpha)g(\alpha) = cf(\alpha)g(\alpha),$$

so $(c - 1)f(\alpha)g(\alpha) = 0$, so $f(\alpha) = 0$, or $g(\alpha) = 0$. Therefore, if we let $W_1$ be the null-space of $f$:

$$W_1 = \{\alpha \in V : f(\alpha) = 0\},$$

and $W_2$ be the nullspace of $g$, then $V = W_1 \cup W_2$. But we know from a previous homework that the union of two subspaces is a subspace exactly when one is contained in the other one. Thus either $W_1 \subset W_2$ or $W_2 \subset W_1$.

In the former case $V = W_1 \cup W_2 = W_2$ so $g = 0$, and in the later case, $V = W_1 \cup W_2 = W_1$, so $f = 0$. 
We know that 
\[ \text{trace}(A + cB) = \text{trace}(A) + c \text{trace}(B), \]
so the set of trace zero matrices is a subspace of \( W \), which we denote by \( W_1 \). Let \( E_{i,j} \) be a matrix whose entries are all zero except the \((i, j)\)-th entry which is equal to 1. Let \( M^i, 1 \leq i \leq n - 1 \) be the matrix whose entries are all zero, except the \((i, i)\)-th entry which is 1 and the \((n, n)\)-th entry which is -1. Then \( E_{i,j}, 1 \leq i, j \leq n, i \neq j \), and \( M^i, 1 \leq i \leq n - 1 \) are all in \( W_1 \). These \((n^2 - n) + (n - 1) = n^2 - n\) matrices span \( W_1 \): If \( A \in W_1 \), \( A \) has the form
\[
\begin{pmatrix}
a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\
a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{n,1} & a_{n,2} & \cdots & -(a_{1,1} + \cdots + a_{n-1,n-1})
\end{pmatrix}
\]
So
\[
A = \sum_{1 \leq i \neq j \leq n} a_{i,j} E_{i,j} + \sum_{i=1}^{n-1} a_{i,i} M^i.
\]
(It is easy to show that these \( n^2 - 1 \) matrices are linearly independent too, so \( \dim W_1 = n^2 - 1 \), but it is not needed here).

Note that if a matrix \( P \) can be written as \( AB - BA \), then the same is true for every scalar multiple of \( P \):
\[ cP = cAB - cBA = (cA)B - B(cA) = A'B - BA' \]
where \( A' = cA \).

Now we show that each matrix \( E_{i,j} \) can be written as \( AB - BA \) for two matrices \( A \) and \( B \) and the same is true for every matrix \( M^i \). Since every matrix of trace zero can be written as a linear combination of the \( E_{i,j} \) and \( M^i \), this shows that every matrix of trace zero can be written as a finite sum
\[
(A_1B_1 - B_1A_1) + \cdots + (A_kB_k - B_kA_k)
\]
for some matrices \( A_i \) and \( B_i \). Which is exactly what the question is asking (well this is one direction, the other direction is trivial: every matrix of the form \( AB - BA \) has trace zero by a previous homework, and the same is true for a sum of the matrices of the form \( AB - BA \)).

Note that for any \( i, j, k, l \), we have
\[
E_{i,k}E_{l,j} = \begin{cases} 
0 & \text{if } k \neq l \\
E_{i,l} & \text{if } k = l
\end{cases}
\]
So if \( i \neq j \),

\[
E^{i,j} = E^{i,i} E^{i,j} - E^{i,j} E^{i,i}.
\]

And if \( 1 \leq i \leq n - 1 \),

\[
M^i = E^{i,n} E^{n,i} - E^{n,i} E^{i,n}.
\]

Page 115, 1) (a) \( g(x_1, x_2) = ax_1 \), (b) \( g(x_1, x_2) = bx_1 - ax_2 \), (c) \( g(x_1, x_2) = (a + b)x_1 + (b - a)x_2 \).

Page 149, 6) If \( j_1, \ldots, j_n \) are distinct, then it is easy to show that \( D \) is \( n \)-linear (I think we proved this in class). Conversely, we assume that \( j_1, \ldots, j_n \) are not distinct and we show that \( D \) is not linear. Assume that \( j_r = j_s \), \( r \neq s \). Let \( j := j_r = j_s \). Assume that \( m \) of the numbers \( j_1, \ldots, j_n \) are equal to \( j \). Then \( m \geq 2 \), and if If we denote the rows of \( A \), by \( \rho_1, \ldots, \rho_n \), then

\[
D(\rho_1, \ldots, c\rho_j, \ldots, \rho_n) = c^m A_{j_1,k_1} A_{j_2,k_2} \cdots A_{j_n,k_n}.
\]

But

\[
cD(\rho_1, \ldots, \rho_j, \ldots, \rho_n) = c A_{j_1,k_1} A_{j_2,k_2} \cdots A_{j_n,k_n}.
\]

If we take \( A \) to be the matrix whose entries are all 1, and if we let \( c \) be a scalar, then the two right hand sides of the above equations are \( c^m \) and \( c \). Since \( m \geq 2 \), we can choose a scalar \( c \) such that \( c^m \neq c \), so \( D \) cannot be linear with respect to the \( j \)-th row.

Page 163, 7) This can be proved using induction. For \( k = 2 \), this is just the special case of equation (5-19) of the book. If we know the equality holds for \( k-1 \), and \( A \) is the given matrix, then we can divide the matrix into 4 blocks:

\[
A_1, \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A_k \end{pmatrix}.
\]

By equation (5-19) determinant of

\[
A = (\det A_1) \cdots (\det A_k).
\]
Page 163, 9) Assume $A$ is $n \times n$. If the determinant rank of $A$ is $r$. Then there is a submatrix $B$ of $A$ consisting of say columns $j_1, \ldots, j_r$ and rows $i_1, \ldots i_r$ of $A$ such that $\det(B) \neq 0$. This implies that the columns of $B$ are linearly independent, in particular, columns $j_1, \ldots, j_r$ of $A$ are linearly independent, so the rank of $A$ is at least $r$. so

\[ \text{rank } (A) \geq \text{determinant rank } (A). \]

On the other hand if $\text{rank}(A)= s$, then there are $s$ linearly independent rows of $A$: call them $i_1, \ldots, i_s$. Let $M$ be the $s \times n$ matrix which is formed by rows $i_1, \ldots, i_s$ of $A$. Then since $M$ has linearly independent rows, the rank of $M$ (which is defined to be the row rank of $M$ and is equal to the column rank of $M$) is $s$, so there are $s$ columns $j_1, \ldots, j_s$ of $M$ which are linearly independent. The matrix which is obtained from rows $j_1, \ldots, j_s$ of $M$ is a submatrix of $A$ with rank $s$, so it is invertible, and its determinant is not equal to zero. So

\[ \text{determinant rank}(A) \geq \text{rank}(A), \]

and the result follows.