Norwegian University of Science and Technology
Institutt for matematiske fag

TMA4145 Linear Methods Fall 2008

Solutions to exercise set 12

1 **a)** Let $f, g \in H'$ and $\lambda \in \mathbb{C}$. Since

$$(f + \lambda g)(x) = f(x) + \lambda g(x)$$

is a continuous linear function on H the set H' is a linear subspace of H.

b) Let $f \in H'$. Since f is continuous at 0 there exists $\delta > 0$ such that if $\|x\| \leq \delta$ then $|f(x)| \leq 1$. Now, for any $y \in H$ we have $\|\delta \frac{y}{\|y\|}\| = \delta$ and therefore

$$|f(\delta \frac{y}{\|y\|})| < 1 \quad \Rightarrow \quad |f(y)| \leq \frac{1}{\delta}\|y\|$$

c) Let $y \in H$. For all $x \in H$ the Cauchy Schwarz inequality gives us

$$|f_y(x)| = |\langle x, y\rangle| \leq \|x\|\|y\|$$

with equality iff $x = \alpha y$ for some $\alpha \in \mathbb{C}$. Therefore

$$\inf\{k \geq 0 | \forall x \in H : |f_y(x)| \leq k\|x\|\} = \|y\|$$

Now given $y, y' \in H$ and $\lambda \in \mathbb{C}$ we calculate

$$f_{y+\lambda y'} = \langle \cdot, y + \lambda y'\rangle = \langle \cdot, y\rangle + \bar{\lambda}\langle \cdot, y'\rangle = f_y + \bar{\lambda} f_{y'}$$

Proving that $y \mapsto f_y$ is conjugate linear. Moreover if $y \mapsto f_y = 0$ then $\langle x, y\rangle = 0$ for all $x \in H$. In particular $\|y\|^2 = \langle y, y\rangle = 0$ implying that $y = 0$. In other words the kernel of $y \mapsto f_y$ is zero and therefore the mapping is injective.

d) For any mapping f between any spaces X and Y we have for all subsets $A \subset Y$ that

$$f^{-1}(A^c) = (f^{-1}(A))^c.$$

A condition for continuity of f is therefore that the preimage of any closed set in Y is closed in X.

Given $f \in H'$ the set

$$\ker f = f^{-1}(\{0\})$$

is therefore closed as $\{0\}$ is a closed set in any normed space. Moreover $\ker f$ is easily checked to be a subspace of H. If f is nonzero then $\ker f \neq H$ and therefore by theorem 4.24 $(\ker f)^\perp \neq \emptyset$. Any $w \in (\ker f)^\perp$ satisfies $f(w) \neq 0$ and $\langle w, x\rangle = 0$ for all $x \in \ker f$.

Using a normalization argument we may pick $y \in (\ker f)^{\perp}$ such that $f(y) = 1$. This allows us to write any $x \in H$ as

$$x = \underbrace{(x - f(x)y)}_{\in \ker f} + \underbrace{f(x)y}_{\in (\ker f)^{\perp}}$$

Taking the inner product with y on both sides yields

$$\langle x, y \rangle = f(x)\langle y, y \rangle = f(x)\|y\|^2$$

And if we set $w_f = \frac{y}{\|y\|^2}$ we obtain

$$\langle x, w_f \rangle = \langle x, \frac{y}{\|y\|^2} \rangle = f(x)\|y\|^2 \frac{1}{\|y\|^2} = f(x)$$

as desired. Uniqueness of w_f follow immediately from theorem 1.5iv.

e) By the previous question any linear functional $f \in H'$ can be written as an inner product $f(x) = \langle x, w_f \rangle$ and hence the mapping $y \mapsto f_y$ is surjective.

The mapping

$$f \to \inf\{k \geq 0 | \forall x \in H : |f(x)| \leq k\|x\|\} = \|w_f\| \tag{1}$$

is a norm derived from an inner product $\langle \cdot, \cdot \rangle_{H'}$ on H' defined by

$$\langle f, g \rangle_{H'} = \langle w_f, w_g \rangle$$

It is trivial to check that $\langle \cdot, \cdot \rangle_{H'}$ defines an inner product on H' and it follows that (1) defines a norm.

2 **a)** If $\{b_1, \ldots, b_n\}$ is an orthogonal set in an inner product space V then $\{\frac{b_1}{\|b_1\|}, \ldots, \frac{b_n}{\|b_n\|}\}$ is an orthonormal set in V. By theorem 4.6 the closest point in $\operatorname{lin}(b_1, \ldots, b_n)$ to any $v \in V$is given by

$$\langle v, \frac{b_1}{\|b_1\|} \rangle \frac{b_1}{\|b_1\|} + \cdots + \langle v, \frac{b_n}{\|b_n\|} \rangle \frac{b_n}{\|b_n\|} = \frac{\langle v, b_1 \rangle}{\langle b_1, b_1 \rangle} b_1 + \ldots + \frac{\langle v, b_n \rangle}{\langle b_n, b_n \rangle} b_n$$

b)

- $$p_1(x) = 1$$
- $$x - \frac{\langle x, p_1 \rangle}{\langle p_1, p_1 \rangle} p_1 = x - \frac{1}{2} \Rightarrow p_2(x) = 2x - 1$$
- $$x^2 - \frac{\langle x^2, p_1 \rangle}{\langle p_1, p_1 \rangle} p_1 - \frac{\langle x^2, p_2 \rangle}{\langle p_2, p_2 \rangle} p_2 = x^2 - x + \frac{1}{6} \Rightarrow p_3(x) = 6x^2 - 6x + 1$$
- $$x^3 - \frac{\langle x^3, p_1 \rangle}{\langle p_1, p_1 \rangle} p_1 - \frac{\langle x^3, p_2 \rangle}{\langle p_2, p_2 \rangle} p_2 - \frac{\langle x^3, p_3 \rangle}{\langle p_3, p_3 \rangle} p_3 = x^3 - \frac{3}{2}x^2 + \frac{3}{5}x - \frac{1}{20}$$
$$\Rightarrow p_4(x) = 20x^3 - 30x^2 + 12x - 1$$

c) As p_1, p_2, p_3, p_4 is an orthogonal basis for Poly_3 we have for any $p \in \mathrm{Poly}_3$:

$$p = \sum_{n=1}^{4} \frac{\langle p, p_n \rangle}{\langle p_n, p_n \rangle} p_n$$

If we observe that $p_n(1) = 1$ for $n = 1, 2, 3, 4$ this gives us

$$p(1) = \sum_{n=1}^{4} \frac{\langle p, p_n \rangle}{\langle p_n, p_n \rangle} = \left\langle p, \sum_{n=1}^{4} \frac{p_n}{\langle p_n, p_n \rangle} \right\rangle$$

Which after a "bit" of calculation yields

$$p(1) = \langle p, 140x^3 - 180x^2 + 60x - 4 \rangle$$

3 **a)** Let E_{ij} be the matrix with 1 at the ij'th entry and 0 at all other entries. Then $E_{ij}^* A = E_{ji} A$ is a matrix whose j'th row is equal to the i'th row of A and zero at all other entries. In particular $\langle A, E_{ij} \rangle = \operatorname{Tr} E_{ij}^* A$ picks out the j'th element of the i'th row of A, i.e. $\langle A, E_{ij} \rangle = a_{ij}$.

In other words $E_{1,1}, E_{1,2}, E_{2,1}, E_{2,2}$ is an orthonormal basis for $\mathrm{Mat}_{2,2}(\mathbb{C})$. By the generalized Pythagoras theorem $\|A\|^2 = \sum |a_{ij}|^2$ so a matrix A depends continuously on its entries. Similarly a polynomial in Poly_2 with the topology derived from

$$\langle f, g \rangle = \int_0^1 f(t)\overline{g(t)}\, dt$$

depends continuously on its coefficients, and therefore the mapping

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto t^2 - (a+d)t + (ad - bc)$$

is continuous.

b) Let $\lambda \in \mathbb{C}$ and define the 2×2-matrix A_λ by

$$A_\lambda = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$$

As $p_{A_\lambda}(t) = (1-t)^2$ and $m_{A_\lambda}(t) | p_{A_\lambda}(t)$ we easily deduce that $m_{A_\lambda}(t) = (1-t)^2$ if $\lambda \neq 0$ and $m_{A_\lambda}(t) = (1-t)$ if $\lambda = 0$. This shows that image of the set of A_λ's, which is a continuous *curve* in $\mathrm{Mat}_{2,2}(\mathbb{C})$, under the mapping $A \mapsto m_A(t)$ has a discontinuity at $\lambda = 0$.

c) Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$,

$$\begin{aligned} A^2 - (a+d)A + (ad-bc)I &= A^2 - \begin{pmatrix} aa + ad - ad + bc & ab + db \\ ac + dc & da + dd - ad + bc \end{pmatrix} \\ &= A^2 - \begin{pmatrix} aa + bc & ab + db \\ ac + dc & dd + bc \end{pmatrix} = A^2 - A^2 = 0 \end{aligned}$$

d) A complex 2×2-matrix has two eigenvalues counted with multiplicity

Case 1: $J_A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$, $P_A(t) = m_A(t) = (\lambda - t)^2$.

Case 2: $J_A = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$, $P_A(t) = (\lambda - t)^2$ and $m_A(t) = (1 - t)$.

Case 3: $J_A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$, $P_A(t) = m_A(t) = (\lambda_1 - t)(\lambda_2 - t)$.

We see that the minimal polynomial divides the characteristic polynomial in all cases.

e) The general conjecture is that $m_A(t)|p_A(t)$ for all $A \in \mathrm{Mat}_{n,n}(\mathbb{C})$.

Suppose that A has n distinct eigenvalues, say $\lambda_1, \ldots, \lambda_n$, with $v_1, \ldots, v_n$ the eigenvectors. The characteristic polynomial of A is $\prod_{j=1}^{n}(t - \lambda_j)$. For any polynomial $q(t)$, $q(A)v_j = q(\lambda_j)v_j$. Thus $c_A(A)v_j = c_A(\lambda_j)v_j = 0$. As the v_j form a basis, we deduce that $c_A(A)v = 0$ for all $v \in V$ and thus $c_A(A) = 0$. Hence $m_A(t)|c_A(t)$.

Now suppose that A is of size $n \times n$ with only one eigenvalue, say λ. The Jordan form of A therefore has λ on the diagonals with some 1s in the diagonal above. The matrix of $A - \lambda I$ just has the 1s. If we square this matrix, we find that the resulting matrix has entries only in the diagonal *above* this one. For example,

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}^2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Continuing this, we see that $(A - \lambda I)^k$ has entries only in the diagonal k above the main diagonal. In particular, $(A - \lambda I)^n = 0$. The actual minimum polynomial may be of lower degree since the entries may disappear earlier; the point is that they have all gone by the nth power. Thus $m_A(t) = (t - \lambda)^k$ for some $k \leq n$. Since $c_A(t) = (t - \lambda)^n$ we therefore have $m_A(t)|c_A(t)$ as required.

For the general case, we essentially mix these two strategies. The Jordan Form says that a matrix looks like a stack of blocks looking like this second case. Therefore we see that the minimum polynomial of each block divides its characteristic polynomial. When combining blocks for *different* eigenvalues, we find that both the characteristic and minimum polynomials multiply—this is essentially the first argument—and so since the result holds for each block, it holds in general.