MATH36001 Solutions: Norms 2015

1.

1. $\|x\|_A = \|Ax\| \geq 0$ since $\|\cdot\|$ is a vector norm and $\|x\|_A = 0 \;\Rightarrow\; Ax = 0 \;\Rightarrow\; x = 0$ because $\mathrm{rank}(A) = n$.
2. $\|\alpha x\|_A = \|\alpha Ax\| = |\alpha|\|Ax\| = |\alpha|\|x\|_A$.
3. $\|x+y\|_A = \|A(x+y)\| = \|Ax + Ay\| \leq \|Ax\| + \|Ay\| = \|x\|_A + \|y\|_A$.

2. The function ν is not a vector norm because $\nu(\alpha x) = |\alpha|\nu(x)$ does not hold for all $\alpha \in \mathbb{C}$, $x \in \mathbb{C}^n$.

3. $\|x\|_\infty = \max\limits_{1\leq i\leq n} |x_i| \leq \left(\sum\limits_{i=1}^n |x_i|^2\right)^{1/2} = \|x\|_2$. Equality is attained, for example, for $x = e_1$, where e_1 is the first column of the identity matrix.

$\|x\|_2^2 = \left(\sum_{i=1}^n |x_i|^2\right) \leq \left(\sum_{i=1}^n |x_i|\right)^2 = \|x\|_1^2$. The equality is attained for $x = e_1$.

Let $e = [1, \cdots, 1]^T \in \mathbb{R}^n$ and let $|x| = [|x_1|, \cdots, |x_n|]^T$. We have

$$\|x\|_1 = \sum_{i=1}^n |x_i| = e^T|x| \leq \|e\|_2 \||x|\|_2 = \sqrt{n}\|x\|_2,$$

using the Cauchy–Schwarz inequality for the last inequality. Equality is attained for $x = e$.

$\|x\|_2 = \left(\sum_{i=1}^n |x_i|^2\right)^{1/2} \leq \sqrt{n}\max_{1\leq i\leq n} |x_i| = \sqrt{n}\|x\|_\infty$. Equality is attained for $x = e$.

4. Let $x \in \mathbb{C}^n$ and let $U \in \mathbb{C}^{n\times n}$ be unitary (i.e. $U^*U = I$). Then $\|Ux\|_2 = (x^*U^*Ux)^{1/2} = (x^*x)^{1/2} = \|x\|_2$. Take $U = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 & -1 \\ 1 & 1 \end{smallmatrix}\right]$ (check that U is unitary) and $x = \left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right]$ so that $Ux = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} -1 \\ 3 \end{smallmatrix}\right]$. Then $\|x\|_1 = 3 \neq \|Ux\|_1 = 4/\sqrt{2}$ and $\|x\|_\infty = 2 \neq \|Ux\|_\infty = 3/\sqrt{2}$.

5. $\|xy^*\|_F^2 = \mathrm{trace}(yx^*xy^*) = (x^*x)\,\mathrm{trace}(yy^*) = \|x\|_2^2 \sum_{i=1}^n |y_i|^2 = \|x\|_2^2\|y\|_2^2$.

$\|(xy^*)v\|_2 = |y^*v|\|x\|_2 \leq \|y\|_2\|v\|_2\|x\|_2$ using the Cauchy–Schwarz inequality. Hence

$$\|xy^*\|_2 = \max_{v\neq 0} \frac{\|(xy^*)v\|_2}{\|v\|_2} \leq \|y\|_2\|x\|_2.$$

Equality is attained for $v = y$.

6. $\|A\|_2^2 = \lambda_{\max}(A^*A) \leq \sum_{i=1}^m \lambda_i(A^*A) = \mathrm{trace}(A^*A) = \|A\|_F^2$.

7. We first show that $\|A\|_1 \le \max_j \sum_{i=1}^m |a_{ij}|$. For all $x \in \mathbb{C}^n$,

$$\begin{aligned}\|Ax\|_1 &= \sum_{i=1}^m |(Ax)_i| = \sum_{i=1}^m \Big|\sum_{j=1}^n a_{ij}x_j\Big| \\ &\le \sum_{i=1}^m \sum_{j=1}^n |a_{ij}||x_j| = \sum_{j=1}^n \sum_{i=1}^m |a_{ij}||x_j| \\ &\le \sum_{j=1}^n \left(\max_k \sum_{i=1}^m |a_{ik}|\right)|x_j| = \left(\max_k \sum_{i=1}^m |a_{ik}|\right)\left(\sum_{j=1}^n |x_j|\right) = \left(\max_k \sum_{i=1}^m |a_{ik}|\right)\|x\|_1.\end{aligned}$$

Therefore $\|Ax\|_1/\|x\|_1 \le \max_k \sum_{i=1}^m |a_{ik}|$ for all $x \ne 0$. From this $\|A\|_1 \le \max_k \sum_{i=1}^m |a_{ik}|$.

To prove equality we must find an $\hat{x}$ for which $\|A\hat{x}\|_1/\|\hat{x}\|_1 = \max_k \sum_{i=1}^m |a_{ik}|$. Suppose that the maximum is attained in the ℓth column of A. Let $\hat{x}$ be the vector with 1 in position ℓ and zeros elsewhere. Then $\|\hat{x}\|_1 = 1$ and $\|A\hat{x}\|_1 = \sum_{i=1}^m |a_{i\ell}|$. Thus $\|A\hat{x}\|_1/\|\hat{x}\|_1 = \sum_{i=1}^m |a_{i\ell}| = \max_k \sum_{i=1}^m |a_{ik}|$.

The expression for $\|\cdot\|_\infty$ is proved in a similar way.

8. If $x \ne 0$ is such that $A^*Ax = \mu^2 x$ with $\mu = \|A\|_2$, then $\mu^2\|x\|_1 = \|A^*Ax\|_1 \le \|A^*\|_1\|A\|_1\|x\|_1 = \|A\|_\infty\|A\|_1\|x\|_1$.

9. Suppose $I - A$ is singular. It follows that $(I-A)x = 0$ for some nonzero x. But then $\|x\| = \|Ax\|$ implies that $\|A\| \ge 1$, a contradiction. Thus $I - A$ is nonsingular. To obtain an expression for its inverse consider the identity $\big(\sum_{k=0}^n A^k\big)(I-A) = I - A^{n+1}$. Since $\|A\| < 1$, it follows that $\lim_{k\to\infty} A^k = 0$ because $\|A^k\| \le \|A\|^k$. Thus $\lim_{n\to\infty}\Big(\sum_{k=0}^n A^k\Big)(I-A) = I$.

10. (i) Since A is Hermitian, the eigendecomposition $A = U\Lambda U^*$ holds with U unitary and $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ real. Then $x^*Ax > 0$ for all nonzero $x \in \mathbb{C}^n \Leftrightarrow x^*U\Lambda U^*x > 0$ for all nonzero $x \in \mathbb{C}^n \Leftrightarrow y^*\Lambda y > 0$ for all $y \in \mathbb{C}^n \Leftrightarrow \lambda_i > 0,\ i = 1, \dots, n$.

(ii) Note that $A^* = A$ so A is Hermitian. Gershgorin's theorem says that the eigenvalues lie in the union of the disks

$$\{\, z : |z-3| \le 1 \,\} \cup \{\, z : |z-3| \le 2 \,\} = \{\, z : |z-3| \le 2 \,\},$$

showing that A cannot have negative or zero eigenvalues. So A is Hermitian positive definite.

11. (i) We prove the result for $n = 3$ but the same technique can be used for arbitrary degree n. Let $n = 3$ so that $C(p) = \begin{bmatrix} -a_2 & -a_1 & -a_0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. Thus $zI - C(p) = \begin{bmatrix} z + a_2 & a_1 & a_0 \\ -1 & z & 0 \\ 0 & -1 & z \end{bmatrix}$.

Adding z^2 times the first column and z times the 2nd column to the last column leaves the determinant unchanged and yields

$$\det\big(zI - C(p)\big) = \det\begin{bmatrix} z+a_2 & a_1 & p(z) \\ -1 & z & 0 \\ 0 & -1 & 0 \end{bmatrix} = p(z)\det\begin{bmatrix} -1 & z \\ 0 & -1 \end{bmatrix} = p(z).$$

(ii) From (i) we conclude that a root $\widetilde{z}$ of $p(z) = 0$ is also an eigenvalue of $C(p)$. It follows that $|\widetilde{z}| \le \|C(p)\|$ for any consistent matrix norm (see Sec. 3 of the handout).

(iii) $|\widetilde{z}| \le \|C(p)\|_1 = \max\{1 + |a_{n-1}|, \cdots, 1 + |a_1|, |a_0|\} \le 1 + \max_{0\le i\le n-1} |a_i|$.

$|\widetilde{z}| \le \|C(p)\|_\infty = \max\{1, |a_{n-1}| + \cdots + |a_0|\} \le 1 + \sum_{i=0}^{n-1} |a_i|$.

Note that Montel's bound is poorer than Cauchy's bound.

12. (i) $M_n e = \mu_n e$ (vector of row sums), $e^T M_n = \mu_n e^T$ (vector of column sums).

(ii) $M_n^2 e = M_n(M_n e) = M_n(\mu_n e) = \mu_n^2 e$, and similarly $e^T M_n^2 = \mu_n^2 e^T$, so M_n^2 has the magic sum μ_n^2 for any n. For example,

$$M_4^2 = \begin{bmatrix} 345 & 257 & 281 & 273 \\ 257 & 313 & 305 & 281 \\ 281 & 305 & 313 & 257 \\ 273 & 281 & 257 & 345 \end{bmatrix}$$

has row and column sums equal to 1156 but it is not a genuine magic square because its entries are not the integers 1 to 16.

(iii) For any magic square matrix, the 1, 2 and ∞ norms are all the same: they equal the magic number, μ_n. The 1 and ∞ norms obviously equal the magic number, so we just have to show that the 2-norm does too.

$$\|M_n\|_2 \ge \frac{\|M_n e\|_2}{\|e\|_2} = \frac{\|\mu_n e\|_2}{\|e\|_2} = \mu_n.$$

From Exercise 8, $\|M_n\|_2 \le \sqrt{\|M_n\|_1\|M_n\|_\infty} = \sqrt{\mu_n \cdot \mu_n} = \mu_n$. Thus $\|M_n\|_2 = \mu_n$.

Since μ_n is an eigenvalue of M_n, $\rho(M_n) \ge \mu_n$ but $\rho(M_n) \le \|M_n\|_1 = \mu_n$ so $\rho(M_n) = \mu_n$.

(iv) Let $B = M_n/\mu_n$. Then B has distinct eigenvalues, one equal to 1 and the others of modulus less than 1. Therefore B is diagonalizable: there is a nonsingular matrix V, whose columns are eigenvectors of B, such that $B = VDV^{-1}$, where $D = \operatorname{diag}(\lambda_i)$, with $\lambda_1 = 1$, say. Hence $B^k = VD^kV^{-1}$ and so

$$\lim_{k\to\infty} B^k = V \lim_{k\to\infty} D^k V^{-1} = V \operatorname{diag}(1,0,0,0)V^{-1} = v_1 w_1^T,$$

where v_1 is the first column of V and w_1^T is the first row of V^{-1}.

But v_1 is an eigenvector of B associated with the eigenvalue $\lambda = 1$ and since $B = M_n/\mu_n$, v_1 is an eigenvector of M_n associated with the eigenvalue $\lambda = \mu_n$. From (i) we can take $v_1 = e$.

We now show that if $v_1 = e$ then $w_1 = \frac{1}{n}e$. First, $e_1^T V^{-1} B = e_1^T D V^{-1} \;\Leftrightarrow\; w_1^T B = w_1^T \;\Leftrightarrow\; w_1^T M_n = \mu_n w_1^T$ so from (i), w_1 must be a multiple of e, say αe. Since $V^{-1}V = I$, $w_1^T v_1 = 1$ which implies that $\alpha = 1/n$. Hence $v_1 w_1^T = ee^T/n$.