Solutions Homework 1 Foundations of Computational Math 1
Fall 2011

Problem 1.1

This problem considers three basic vector norms: $\|.\|_1$,$\|.\|_2$,$\|.\|_\infty$.

1.1.a Prove that $\|.\|_1$ is a vector norm.

1.1.b Prove that $\|.\|_\infty$ is a vector norm.

1.1.c Consider $\|.\|_2$.

- i Show that $\|.\|_2$ is definite.
- ii Show that $\|.\|_2$ is homogeneous.
- iii Show that for $\|.\|_2$ the triangle inequality follows from the Cauchy inequality $|x^H y| \leq \|x\|_2 \|y\|_2$.
- iv Assume you have two vectors x and y such that $\|x\|_2 = \|y\|_2 = 1$ and $x^H y = |x^H y|$, prove the Cauchy inequality holds for x and y.
- v Assume you have two arbitrary vectors $\tilde{x}$ and $\tilde{y}$. Show that there exists x and y that satisfy the conditions of part (iv) and $\tilde{x} = \alpha x$ and $\tilde{y} = \beta y$ where α and β are scalars.
- vi Show the Cauchy inequality holds for two arbitrary vectors $\tilde{x}$ and $\tilde{y}$.

Solution: We consider each norm in turn.

One Norm: If $x \neq 0$ then there exists an element $\xi_j \neq 0$. Therefore,

$$\begin{aligned} \|x\|_1 &= \Sigma_i |\xi_i| \\ &\geq |\xi_j| \\ &\geq 0 \end{aligned}$$

and $x = 0$ implies that $\|x\|_1 = 0$. Since all terms in the sum are nonnegative the only way the sum can be 0 is if all terms are 0 which implies all $\xi_i = 0$. It follows that $\|x\|_1 = 0$ implies $x = 0$. Therefore $\|x\|_1$ is definite.

We have

$$\begin{aligned}\|\alpha x\|_1 &= \Sigma_i|\alpha\xi_i| \\ &= \Sigma_i|\alpha||\xi_i| \\ &= |\alpha|\Sigma_i|\xi_i| \\ &= |\alpha|\|x\|_1\end{aligned}$$

and therefore $\|x\|_1$ is homogeneous.

We have, given the triangle inequality for magnitude on $\mathbb{R}$ and $\mathbb{C}$,

$$\begin{aligned}\|x+y\|_1 &= \Sigma_i|\xi_i+\eta_i| \\ &\leq \Sigma_i(|\xi_i|+|\eta_i|) \\ &= \Sigma_i|\xi_i|+\Sigma_i|\eta_i| \\ &= \|x\|_1+\|y\|_1\end{aligned}$$

and therefore $\|x\|_1$ satisfies the triangle inequality.

Max Norm: If $x \neq 0$ then there exists an element $\xi_j \neq 0$. Therefore,

$$\begin{aligned}\|x\|_\infty &= \max_i |\xi_i| \\ &\geq |\xi_j| \\ &\geq 0\end{aligned}$$

Therefore $\|x\|_\infty$ is definite.

We have

$$\begin{aligned}\|\alpha x\|_\infty &= \max_i |\alpha\xi_i| \\ &= \max_i(|\alpha||\xi_i|) \\ &= |\alpha| \max_i |\xi_i| \\ &= |\alpha|\|x\|_\infty\end{aligned}$$

and therefore $\|x\|_\infty$ is homogeneous.

We have

$$\begin{aligned}\|x+y\|_\infty &= \max_i |\xi_i+\eta_i| \\ &\leq \max_i(|\xi_i|+|\eta_i|) \\ &\leq \max_i |\xi_i| + \max_i |\eta_i| \\ &= \|x\|_\infty+\|y\|_\infty\end{aligned}$$

and therefore $\|x\|_\infty$ satisfies the triangle inequality.

Two Norm: If $x \neq 0$ then there exists an element $\xi_j \neq 0$. Therefore,

$$\begin{aligned} \|x\|_2^2 &= \Sigma_i |\xi_i|^2 \\ &\geq |\xi_j|^2 \\ &> 0 \end{aligned}$$

Therefore $\|x\|_2$ is definite.

We have

$$\begin{aligned} \|\alpha x\|_2^2 &= \Sigma_i |\alpha \xi_i|^2 \\ &= \Sigma_i (|\alpha|^2 |\xi_i|^2) \\ &= |\alpha|^2 \Sigma_i |\xi_i|^2 \\ &= |\alpha|^2 \|x\|_2 \end{aligned}$$

and therefore $\|x\|_2$ is homogeneous.

The triangle inequality follows from the Cauchy inequality for the two norm:

$$|x^H y| \leq \|x\|_2 \|y\|_2$$

as follows

$$\begin{aligned} \|x+y\|_2^2 &= x^H x + y^H y + 2\mathcal{R}e(x^H y) \\ &= \|x\|_2^2 + \|y\|_2^2 + 2\mathcal{R}e(x^H y) \\ &\leq \|x\|_2^2 + \|y\|_2^2 + 2|x^H y| \\ &\leq \|x\|_2^2 + \|y\|_2^2 + 2\|x\|_2 \|y\|_2 \\ &= (\|x\|_2 + \|y\|_2)^2. \end{aligned}$$

So the true problem is to prove the Cauchy inequality. To do so assume that we have two vectors x and y such that $\|x\|_2 = \|y\|_2 = 1$ and $x^H y = |x^H y|$. For any two such vectors we have

$$\begin{aligned} \|x-y\|_2^2 &= (x-y)^H (x-y) \\ &= 2 - 2x^H y \\ &\geq 0 \end{aligned}$$

Therefore $|x^H y| \leq 1 = \|x\|_2 \|y\|_2$.

To generalize to any two nonzero vectors $\tilde{x}$ and $\tilde{y}$ note that there must exist complex scalars α and β such that $\tilde{x} = \alpha x$ and $\tilde{y} = \beta y$ where x and y satisfy the conditions above (see Lemma below). We have

$$\begin{aligned} |\tilde{x}^H \tilde{y}| &= |\alpha x^H y \beta| \\ &= |\alpha\beta| |x^H y| \\ &\leq |\alpha\beta| \|x\|_2 \|y\|_2 \\ &= |\alpha| \|x\|_2 |\beta| \|y\|_2 \\ &= \|\alpha x\|_2 \|\beta y\|_2 \\ &= \|\tilde{x}\|_2 \|\tilde{y}\|_2 \end{aligned}$$

Lemma. *Let $\tilde{x} \in \mathbb{C}^n$ and $\tilde{y} \in \mathbb{C}^n$ be such that $\tilde{x} \neq 0$ and $\tilde{y} \neq 0$. There exists $\alpha \in \mathbb{C}$, $\beta \in \mathbb{C}$, $x \in \mathbb{C}^n$ and $y \in \mathbb{C}^n$ such that*

$$\|x\|_2 = \|y\|_2 = 1 \qquad |x^H y| = x^H y$$
$$\tilde{x} = \alpha x \qquad \tilde{y} = \beta y$$

Proof. Suppose $\tilde{x}^H \tilde{y} = \gamma e^{i\phi}$ where $\gamma \in \mathbb{R}$ and $\gamma > 0$. Let $\phi_1 \in \mathbb{R}$ and $\phi_2 \in \mathbb{R}$ be such that $\phi = \phi_1 + \phi_2$. The scalars α and β can be set as follows:

$$\alpha = \|\tilde{x}\| e^{-i\phi_1} \qquad \beta = \|\tilde{y}\| e^{i\phi_2}$$

Taking $\tilde{x} = \alpha x$ and $\tilde{y} = \beta y$ implies that

$$\|x\|_2^2 = x^H x = \frac{1}{|\alpha|^2} \tilde{x}^H \tilde{x} = \frac{1}{\|\tilde{x}\|_2^2} \tilde{x}^H \tilde{x} = 1$$

$$\|y\|_2^2 = y^H y = \frac{1}{|\beta|^2} \tilde{y}^H \tilde{y} = \frac{1}{\|\tilde{y}\|_2^2} \tilde{y}^H \tilde{y} = 1$$

So $\|x\|_2 = \|y\|_2 = 1$ and we have that

$$\gamma e^{i\phi} = \tilde{x}^H \tilde{y} = \bar{\alpha}\beta \|\tilde{y}\|_2 \|\tilde{x}\|_2 x^H y = e^{i\phi_1} e^{i\phi_2} \|\tilde{y}\|_2 \|\tilde{x}\|_2 x^H y = e^{i(\phi_1+\phi_2)} \|\tilde{y}\|_2 \|\tilde{x}\|_2 x^H y = e^{i\phi} \|\tilde{y}\|_2 \|\tilde{x}\|_2 x^H y$$

$$\therefore x^H y = \frac{\gamma}{\|\tilde{y}\|_2 \|\tilde{x}\|_2}$$

Since γ, $\|\tilde{y}\|$ and $\|\tilde{x}\|$ are real and positive it follows that $x^H y$ is also real and positive. Therefore, $x^H y = |x^H y|$ as desired. □

Problem 1.2

What is the unit ball in $\mathbb{R}^2$ for each of the vector norms: $\|.\|_1$,$\|.\|_2$,$\|.\|_\infty$?
Solution:

The unit balls in $\mathbb{R}^2$ follow immediately from the equations $\|x\| = 1$ for each norm.

- $\|x\|_2 = 1 = \|x\|_2^2 = \xi_1^2 + \xi_2^2 \rightarrow$ a circle with radius 1 centered at the origin.
- $\|x\|_\infty = 1 = max(|\xi_1|, |\xi_2|) \rightarrow$ a square with corners $(1,1)$, $(-1,1)$, $(-1,-1)$, $(1,-1)$.
- $\|x\|_1 = 1 = |\xi_1|+|\xi_2| \rightarrow$ a diamond with sides of length 1 and corners $(1,0)$, $(0,1)$, $(-1,0)$, $(0,-1)$.

Problem 1.3

Consider the matrices

$$B_1 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} \quad B_2 = \begin{pmatrix} 0 & 2 \\ 0 & 2 \\ -1 & 1 \end{pmatrix}$$

1.3.a. Show that they have the same range space.

1.3.b. We have $x = B_1 c_1 = B_2 c_2$ for all x in the range space. Determine the relationship between c_1 and c_2 and express it as a linear transformation.

Solution: The linear independence of the columns of B_1 and B_2 is clear from the zero/nonzero structure. So $\mathcal{R}(B_1)$ and $\mathcal{R}(B_2)$ are each subspaces of $\mathbb{R}^3$ with dimension 2 and **the columns of the respective matrices are bases for the corresponding spaces**. To see that the spaces are the same it is sufficient to show that the columns of B_1 can be written as linear combinations of the columns of B_2 and vice versa. This implies that the columns of each matrix are **also a basis for the range space of the other matrix**.

Specifically, we have

$$\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = -\frac{1}{2} \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$$

So the columns of B_1 are 2 linearly independent vectors in $\mathcal{R}(B_2)$. Since the dimension of $\mathcal{R}(B_2)$ is 2, the columns of B_1 are a basis for $\mathcal{R}(B_2)$ and therefore $\mathcal{R}(B_2) = \mathcal{R}(B_1)$.

Alternatively, We have

$$\begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

So the columns of B_2 are 2 linearly independent vectors in $\mathcal{R}(B_1)$. Since the dimension of $\mathcal{R}(B_1)$ is 2, the columns of B_2 are a basis for $\mathcal{R}(B_1)$ and therefore $\mathcal{R}(B_2) = \mathcal{R}(B_1)$.

These relationships can all be written in terms of a linear transformation relating B_1 and B_2 to reach the same conclusions. To see this note that

$$B_1 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 0 & 2 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = B_2 M$$

$$B_2 = \begin{pmatrix} 0 & 2 \\ 0 & 2 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = B_1 N$$

Therefore, any linear combination of the columns of B_1 is in $\mathcal{S}$ and can be written as a linear combination of the columns of B_1 and vice versa. The columns of M are easily seen to be linearly independent as are the columns of N. It is also easily verified that $N = M^{-1}$.

We then have for any $x \in \mathcal{S} \subset \mathbb{R}^3$ we have unique $c_1 \in \mathbb{R}^2$ and $c_2 \in \mathbb{R}^2$

$$x = B_1 c_1 = (B_2 M) c_1 = B_2 (M c_1) = B_2 c_2$$

and since M is nonsingular it relates all such c_1 and c_2, i.e. it is an invertible map from $\mathbb{R}^2$ to itself. The conclusion $\mathcal{R}(B_2) = \mathcal{R}(B_1)$ follows.

Problem 1.4

Let $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear function, i.e.,

$$F(\alpha x + \beta y) = \alpha F(x) + \beta F(y)$$

.

1.4.a. Suppose you are given a routine that returns $F(x)$ given any $x \in \mathbb{R}^n$. How would you use this routine to determine a matrix $A \in \mathbb{R}^{m \times n}$ such that $F(x) = Ax$ for all $x \in \mathbb{R}^n$?

1.4.b. Show A is unique.

Solution: For the standard basis $e_1, \cdots, e_n$ of $\mathbb{R}^n$ we have uniquely $\forall x \in \mathbb{R}^n$

$$x = \sum_{i=1}^{n} e_i \xi_i.$$

By linearity we have

$$F(x) = F(\sum_{i=1}^{n} e_i \xi_i) = \sum_{i=1}^{n} F(e_i) \xi_i = \sum_{i=1}^{n} a_i \xi_i = Ax$$

where $F(e_i) = a_i = Ae_i$ determines the i-th column of A. So evaluating F on the n standard basis vectors yields A.

Suppose we are given two matrices A and B that define $F(\cdot)$. We have by definition

$$\forall x \in \mathbb{R}^n \;\; y = Ax = Bx$$
$$\therefore Ae_i = Be_i \;\; 1 \le i \le n$$
$$\therefore A = B$$

Problem 1.5

Consider the matrix

$$L = \begin{pmatrix} \lambda_{11} & 0 & 0 & 0 \\ \lambda_{21} & \lambda_{22} & 0 & 0 \\ \lambda_{31} & \lambda_{32} & \lambda_{33} & 0 \\ \lambda_{41} & \lambda_{42} & \lambda_{43} & \lambda_{44} \end{pmatrix}$$

Suppose that $\lambda_{11} \neq 0$, $\lambda_{33} \neq 0$, $\lambda_{44} \neq 0$ but $\lambda_{22} = 0$.

1.5.a. Show that L is singular.

1.5.b. Determine a basis for the nullspace $\mathcal{N}(L)$.

Solution: To show that L is singular and find the nullspace we must determine the structure of the $x \neq 0$ such that $Lx = 0$. Imposing the λ_{ii} constraints we have, $\lambda_{11} \neq 0 \rightarrow \xi_1 = 0$ and therefore,

$$\begin{pmatrix} \lambda_{11} & 0 & 0 & 0 \\ \lambda_{21} & 0 & 0 & 0 \\ \lambda_{31} & \lambda_{32} & \lambda_{33} & 0 \\ \lambda_{41} & \lambda_{42} & \lambda_{43} & \lambda_{44} \end{pmatrix} \begin{pmatrix} 0 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

where ξ_2 is arbitrary and

$$\begin{pmatrix} \lambda_{33} & 0 \\ \lambda_{43} & \lambda_{44} \end{pmatrix} \begin{pmatrix} \xi_3 \\ \xi_4 \end{pmatrix} = \begin{pmatrix} -\lambda_{32} \\ -\lambda_{42} \end{pmatrix} \xi_2.$$

Since $\lambda_{33} \neq 0$ and $\lambda_{44} \neq 0$ the 2×2 matrix is nonsingular and therefore, ξ_3 and ξ_4 are uniquely determined given a particular value for ξ_2, i.e., there are no further degrees of freedom in the null space vectors. It follows that the dimension of $\mathcal{N}(L)$ is 1. So

$$\mathcal{N}(L) = span[\begin{pmatrix} 0 \\ 1 \\ \nu_2 \\ \nu_3 \end{pmatrix}] \text{ where } \begin{pmatrix} \lambda_{33} & 0 \\ \lambda_{43} & \lambda_{44} \end{pmatrix} \begin{pmatrix} \nu_2 \\ \nu_3 \end{pmatrix} = \begin{pmatrix} -\lambda_{32} \\ -\lambda_{42} \end{pmatrix}$$

Problem 1.6

1.6.a Let $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times n}$ be nonsingular matrices. Show $(AB)^{-1} = B^{-1}A^{-1}$.

1.6.b Suppose $A \in \mathbb{R}^{m\times n}$ with $m > n$ and let $M \in \mathbb{R}^{n\times n}$ be a nonsingular square matrix. Show that $\mathcal{R}(A) = \mathcal{R}(AM)$ where $\mathcal{R}()$ denotes the range of a matrix.

Solution:

By assumption, the $n \times n$ nonsingular matrices A^{-1} and B^{-1} exist, are unique and $AA^{-1} = BB^{-1} = A^{-1}A = B^{-1}B = I$. Let $M = B^{-1}A^{-1}$.

We have

$$\begin{aligned}
(AB)M &= ABB^{-1}A^{-1} \\
&= A(BB^{-1})A^{-1} \\
&= AIA^{-1} \\
&= AA^{-1} \\
&= I \\
M(AB) &= B^{-1}A^{-1}AB \\
&= B^{-1}(A^{-1}A)B \\
&= B^{-1}IB \\
&= B^{-1}B \\
&= I
\end{aligned}$$

To see that M is unique, suppose there is another matrix $Q \neq M$ such that $ABQ = I$. We have

$$Q = IQ = (MAB)Q = M(ABQ) = M$$

Now suppose that $Q \neq M$ such that $QAB = I$. We have

$$Q = QI = Q(ABM) = (QAB)M = M$$

(Strictly speaking you need only prove one of these to show uniqueness.)

We must show $\mathcal{R}(A) \subseteq \mathcal{R}(AM)$ and $\mathcal{R}(A) \supseteq \mathcal{R}(AM)$.

We have $y \in \mathcal{R}(A) \to \exists x \in \mathbb{R}^n$ such that $y = Ax$. M nonsingular implies that $\forall x \in \mathbb{R}^n \exists c \in \mathbb{R}^n$ such that $x = Mc$. Therefore, $y = AMc$ and $y \in \mathcal{R}(AM)$.

We have $y \in \mathcal{R}(AM) \to \exists x \in \mathbb{R}^n$ such that $y = AMx$. Also $M \in \mathbb{R}^{n\times n} \to b = Mx \in \mathbb{R}^n$. Therefore, $y = AMx = Ab \to y \in \mathcal{R}(A)$.

Problem 1.7

Let $y \in \mathbb{R}^m$ and $\|y\|$ be any vector norm defined on $\mathbb{R}^m$. Let $x \in \mathbb{R}^n$ and A be an $m \times n$ matrix with $m > n$.

1.7.a. Show that the function $f(x) = \|Ax\|$ is a vector norm on $\mathbb{R}^n$ if and only if A has full column rank, i.e., $rank(A) = n$.

1.7.b. Suppose we choose $f(x)$ from part (1.7.a) to be $f(x) = \|Ax\|_2$. What condition on A guarantees that $f(x) = \|x\|_2$ for any vector $x \in \mathbb{R}^n$?

Solution:

This question essentially asks when can we embed the vector space $\mathbb{R}^n$ in $\mathbb{R}^m$ in order to define a norm on $\mathbb{R}^n$.

$\rightarrow$

Let $y \in \mathbb{R}^m$ and $g(y) = \|y\|$ be a vector norm on $\mathbb{R}^m$. We know by the definiteness of norms that if $g(y) = 0$ only when $y = 0$. So, since $y = Ax$, we must consider what x can lead to $y = 0$. By assumption, $f(x)$ is a vector norm for $\mathbb{R}^n$. We know by the definiteness of norms that $f(x) = 0$ only when $x = 0$.

Now assume that $rank(A) < n$. This means that there exists an $x \neq 0$ such that $Ax = 0$. Therefore we have $\exists x \neq 0$ such that $f(x) = f(Ax) = f(0) = 0$. This contradicts the assumption that $f(x)$ is a vector norm on $\mathbb{R}^n$. Therefore, $rank(A) = n$ must hold if $f(x)$ is a vector norm on $\mathbb{R}^n$.

$\leftarrow$

Now suppose $rank(A) = n$. Since $g(y)$ is a norm on $\mathbb{R}^n$ we need only check that $f(x)$ satisfies the properties of a norm by writing it in terms of $g(y)$.

Since A is full rank, $x \neq 0 \rightarrow y = Ax \neq 0$. Therefore,

$$\begin{aligned} f(x) &= g(Ax) \\ &= g(y) \\ &= \|y\| \\ &\neq 0 \end{aligned}$$

and f is definite.

Let $x \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$.

$$\begin{array}{rcccc} f(\alpha x) & = & g(A(\alpha x)) & = & g(\alpha Ax) \\ & = & g(\alpha y) & = & \|\alpha y\| \\ & = & |\alpha|\|y\| & = & |\alpha| g(y) \\ & = & |\alpha| g(Ax) & = & |\alpha| f(x) \end{array}$$

Therefore $f(x)$ is homogeneous.

Let $x_1 \in \mathbb{R}^n$ and $x_2 \in \mathbb{R}^n$.

$$\begin{array}{rcccc} f(x_1 + x_2) & = & g(A(x_1 + x_2)) & = & g(Ax_1 + Ax_2) \\ & = & g(y_1 + y_2) & = & \|y_1 + y_2\| \\ & \leq & \|y_1\| + \|y_2\| & = & g(y_1) + g(y_2) \\ & = & g(Ax_1) + g(Ax_2) & = & f(x_1) + f(x_2) \end{array}$$

For the second part of the question, if the matrix $A \in \mathbb{R}^{m \times n}$ is an isometry, i.e., it has orthonormal columns, then $A^T A = I_n$. We therefore have

$$\begin{aligned} f(x)^2 &= \|Ax\|_2^2 \\ &= x^T A^T A x \\ &= x^T x \\ &= \|x\|_2^2 \end{aligned}$$

and $f(x) = \|x\|_2$ on $\mathbb{R}^n$ as desired.