Basic Concepts in Linear Algebra

Grady B. Wright

Department of Mathematics
Boise State University

February 2, 2015

Numerical Linear Algebra

- Linear systems of equations occur in almost every area of the applied science, engineering, and mathematics.
- Hence, numerical linear algebra is one of the pillars of computational mathematics.

Linear systems of equations

A linear system of m equations and n unknowns can be expressed in the following general form:

$$\begin{array}{ccccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & a_{13}x_3 & + & \cdots & + & a_{1n}x_n & = & b_1, \\ a_{21}x_1 & + & a_{22}x_2 & + & a_{23}x_3 & + & \cdots & + & a_{2n}x_n & = & b_2, \\ a_{31}x_1 & + & a_{32}x_2 & + & a_{33}x_3 & + & \cdots & + & a_{3n}x_n & = & b_3, \\ \vdots & & \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & a_{m3}x_3 & + & \cdots & + & a_{mn}x_n & = & b_m. \end{array} \tag{1}$$

Here a_{ij} are the coefficients of the systems, b_i are the right hand sides (RHS), and x_j are the unknown values that must be determined. a_{ij} and b_i will be given by the problem.

Linear systems of equations

Linear systems can be classified into the following three types:

1. **Square linear system**: If the number of equations equals the number of unknowns (i.e. $m = n$).
2. **Overdetermined system**: If the number of equations is greater than the number of unknowns (i.e. $m > n$).
3. **Underdetermined system**: If the number equations is less than the number of unknowns (i.e. $m < n$).

Matrices and vectors

A convenient notation to describe a linear system of equations is in terms of matrices and vectors.

Matrices

- A matrix is just a table of numbers containing m rows and n columns and can be expressed as:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}.$$

- We typically use capital letters to denote matrices.
- We write $A \in \mathbb{R}^{m\times n}$ to denote a matrix with m rows and n columns.
- A common shorthand notation for a matrix is $A = \{a_{ij}\}$, where the values for i and j are understood from the problem.

Matrices

- A matrix is just a table of numbers containing m rows and n columns and can be expressed as:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}.$$

- We typically use capital letters to denote matrices.
- We write $A \in \mathbb{R}^{m\times n}$ to denote a matrix with m rows and n columns.
- A common shorthand notation for a matrix is $A = \{a_{ij}\}$, where the values for i and j are understood from the problem.

Matrices

- A matrix is just a table of numbers containing m rows and n columns and can be expressed as:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}.$$

- We typically use capital letters to denote matrices.
- We write $A \in \mathbb{R}^{m\times n}$ to denote a matrix with m rows and n columns.
- A common shorthand notation for a matrix is $A = \{a_{ij}\}$, where the values for i and j are understood from the problem.

Matrices

- A matrix is just a table of numbers containing m rows and n columns and can be expressed as:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}.$$

- We typically use capital letters to denote matrices.
- We write $A \in \mathbb{R}^{m \times n}$ to denote a matrix with m rows and n columns.
- A common shorthand notation for a matrix is $A = \{a_{ij}\}$, where the values for i and j are understood from the problem.

Vectors

If the matrix only has one row or column then it is called a **vector**.

- A **column vector** with n entries can be expressed as

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}.$$

- A **row vector** and can be expressed as

$$\mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 & \cdots & x_n \end{bmatrix}.$$

- We typically use bold lower-case letters to denote vectors.
- A column vector with n real entries is denoted by $\mathbf{x} \in \mathbb{R}^n$, while a row vector is denoted by $\mathbf{x} \in \mathbb{R}^{1\times n}$.

Vectors

If the matrix only has one row or column then it is called a **vector.**

- A **column vector** with n entries can be expressed as

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}.$$

- A **row vector** and can be expressed as

$$\mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 & \cdots & x_n \end{bmatrix}.$$

- We typically use bold lower-case letters to denote vectors.
- A column vector with n real entries is denoted by $\mathbf{x} \in \mathbb{R}^n$, while a row vector is denoted by $\mathbf{x} \in \mathbb{R}^{1\times n}$.

Vectors

If the matrix only has one row or column then it is called a **vector.**

- A **column vector** with n entries can be expressed as

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}.$$

- A **row vector** and can be expressed as

$$\mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 & \cdots & x_n \end{bmatrix}.$$

- We typically use bold lower-case letters to denote vectors.
- A column vector with n real entries is denoted by $\mathbf{x} \in \mathbb{R}^n$, while a row vector is denoted by $\mathbf{x} \in \mathbb{R}^{1\times n}$.

Vectors

If the matrix only has one row or column then it is called a **vector.**

- A **column vector** with n entries can be expressed as

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}.$$

- A **row vector** and can be expressed as

$$\mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 & \cdots & x_n \end{bmatrix}.$$

- We typically use bold lower-case letters to denote vectors.
- A column vector with n real entries is denoted by $\mathbf{x} \in \mathbb{R}^n$, while a row vector is denoted by $\mathbf{x} \in \mathbb{R}^{1\times n}$.

Vectors

If the matrix only has one row or column then it is called a **vector.**

- A **column vector** with n entries can be expressed as

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}.$$

- A **row vector** and can be expressed as

$$\mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 & \cdots & x_n \end{bmatrix}.$$

- We typically use bold lower-case letters to denote vectors.
- A column vector with n real entries is denoted by $\mathbf{x} \in \mathbb{R}^n$, while a row vector is denoted by $\mathbf{x} \in \mathbb{R}^{1\times n}$.

Matrix & vector operations

Matrix & vector operations

Matrix & vector operations: Transpose

Let $A \in \mathbb{R}^{m\times n}$ with entries

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix},$$

then the **transpose** of A switches the columns of A with the rows, i.e.

$$A^T = \begin{bmatrix} a_{11} & a_{21} & a_{31} & \cdots & a_{m1} \\ a_{12} & a_{22} & a_{32} & \cdots & a_{m2} \\ a_{13} & a_{23} & a_{33} & \cdots & a_{m3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & a_{3n} & \cdots & a_{mn} \end{bmatrix}.$$

Note that $A^T \in \mathbb{R}^{n\times m}$ and that $(A^T)^T = A$.

Matrix & vector operations: Transpose

The transpose can also be applied to vectors. In this case if $\mathbf{x}$ is a (column) vector then $\mathbf{x}^T$ is a row vector:

$$\text{if} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} \quad \text{then} \quad \mathbf{x}^T = \begin{bmatrix} x_1 & x_2 & x_3 & \cdots & x_n \end{bmatrix}.$$

Similarly if $\mathbf{x}$ is row vector then $\mathbf{x}^T$ is a column vector.

Matrix & vector operations: Matrix addition

Let $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{m\times n}$ then the sum of A and B is given by

$$A + B = \left\{ a_{ij} + b_{ij} \right\}.$$

This is just the sum of the corresponding entries of the elements of A and B.

For this sum to make sense A and B must be the same size.

Matrix & vector operations: scalar multiplication

Let α be a real number and $A \in \mathbb{R}^{m \times n}$ then the product of α and A is given by

$$\alpha A = \left\{ \alpha\, a_{ij} \right\}.$$

Note that this is just α times each entry of A.

Matrix & vector operations: Vector-vector products

There are two types of vector-vector products that arise quite frequently. These can be derived from the definition for matrix-matrix products (discussed later), but it is worth stating them separately.

- Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ then the **inner product** or **dot product** of $\mathbf{x}$ and $\mathbf{y}$ is

$$\mathbf{x}^T\mathbf{y} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n = \sum_{j=1}^{n} x_j y_j.$$

  Note that the inner product is a single number. The inner product is sometimes denoted by $\mathbf{x} \cdot \mathbf{y}$.

Matrix & vector operations: Vector-vector products

There are two types of vector-vector products that arise quite frequently. These can be derived from the definition for matrix-matrix products (discussed later), but it is worth stating them separately.

- Let $\mathbf{x} \in \mathbb{R}^m$ and $\mathbf{y} \in \mathbb{R}^n$ then the **outer product** of $\mathbf{x}$ with $\mathbf{y}$ is

$$\mathbf{x}\mathbf{y}^T = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix} = \begin{bmatrix} x_1 y_1 & x_1 y_2 & \cdots & x_1 y_n \\ x_2 y_1 & x_2 y_2 & \cdots & x_2 y_n \\ \vdots & \vdots & \ddots & \vdots \\ x_m y_1 & x_m y_2 & \cdots & x_m y_n \end{bmatrix}$$

Note that the outer product is a matrix of size m-by-m.

Matrix & vector operations: Matrix-vector products

Let $A \in \mathbb{R}^{m\times n}$ and $\mathbf{x} \in \mathbb{R}^n$ then the product of A and $\mathbf{x}$ is given by

$$A\mathbf{x} = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \\ \vdots \\ a_{m2} \end{bmatrix} + x_3 \begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \\ \vdots \\ a_{m3} \end{bmatrix} + \ldots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ a_{3n} \\ \vdots \\ a_{mn} \end{bmatrix} \tag{2}$$

Thus, the product $A\mathbf{x}$ is a linear combination of the columns of A.

Matrix & vector operations: Matrix-vector products

Important observations regarding the matrix-vector product $A\mathbf{x}$:

- The only way for this product to make sense is if A has the same number of columns as $\mathbf{x}$ does rows.
- $A\mathbf{x} \in \mathbb{R}^m$, i.e. the product is a column vector containing m entries.
- If we let $\mathbf{b} = A\mathbf{x}$ then we can alternatively express the ith entry of $\mathbf{b}$ as

$$b_i = \sum_{j=1}^{n} a_{ij} x_j \,, \quad i = 1, \ldots, m.$$

  This illustrates that b_i is just the inner product of the ith row of A with the vector $\mathbf{x}$.
- In general, computing $A\mathbf{x}$ using the above formulas requires mn multiplications and $m(n-1)$ additions.

Matrix & vector operations: Matrix-matrix products

Let $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{n\times p}$, and let B have columns

$$B = \left[\begin{array}{c|c|c|c} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n \end{array}\right].$$

The matrix-matrix product $C = AB$ is given as

$$C = \left[\begin{array}{c|c|c|c} A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_n \end{array}\right].$$

This shows the kth column of the product AB is a linear combination of the columns of A with the coefficients in the linear combinations being determined by entries in the kth column of B.

Matrix & vector operations: Matrix-matrix products

Important observations regarding the matrix-matrix product AB, $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times p}$:

- Number rows of A must equal number columns B.
- $AB \in \mathbb{R}^{m \times p}$, i.e. the product is a matrix containing m rows and p columns.
- In general, $AB \neq BA$, i.e. the product does not commute.

Matrix & vector operations: Matrix-matrix products

Important observations regarding the matrix-matrix product AB, $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times p}$:

- We can express each entry of C as

$$c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}\,, \quad i = 1, \ldots, m,\;\; k = 1, \ldots, p.$$

  So c_{ik} is just the inner product of the ith row of A with the kth column of B.
- Computing AB using the above formulas requires mnp multiplications and $m(n-1)p$ additions.
- The transpose of the product AB satisfies: $(AB)^T = B^T A^T$.

Linear systems in matrix-vector notation

Recall that we can express a linear system of equations with m equations and n unknowns as

$$\begin{array}{ccccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & a_{13}x_3 & + & \cdots & + & a_{1n}x_n & = & b_1, \\ a_{21}x_1 & + & a_{22}x_2 & + & a_{23}x_3 & + & \cdots & + & a_{2n}x_n & = & b_2, \\ a_{31}x_1 & + & a_{32}x_2 & + & a_{33}x_3 & + & \cdots & + & a_{3n}x_n & = & b_3, \\ \vdots & & \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & a_{m3}x_3 & + & \cdots & + & a_{mn}x_n & = & b_m. \end{array} \tag{3}$$

We can express this linear system in matrix-vector notation using the previous definitions.

Linear systems in matrix-vector notation

Let $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$, and $A \in \mathbb{R}^{m\times n}$, then the linear system is given as $A\mathbf{x} = \mathbf{b}$, or

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{bmatrix}}_{\mathbf{b}}.$$

Linear systems: solvability

Recall that $A\mathbf{x}$ is a linear combination of the columns of A:

$$A\mathbf{x} = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \\ \vdots \\ a_{m2} \end{bmatrix} + x_3 \begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \\ \vdots \\ a_{m3} \end{bmatrix} + \ldots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ a_{3n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

Thus, the only way there will be a solution to $A\mathbf{x} = \mathbf{b}$ is if $\mathbf{b}$ can be written as a linear combination of the columns of A.

Linear systems: solvability

There are three possibilities for the linear system $A\mathbf{x} = \mathbf{b}$:

1. There are an infinite number of solutions that satisfy $A\mathbf{x} = b$.
   An infinite number of ways to linearly combine the columns of A to equal $\mathbf{b}$.
2. There is one unique solution to the linear system.
   Only one way to linearly combine the columns of A to equal $\mathbf{b}$.
3. There is no solution to the linear system.
   There is no way to linearly combine the columns of A to equal $\mathbf{b}$.

Special types of matrices: Diagonal matrix

A diagonal matrix is an n-by-n square matrix with zeros on in every entry except possibly the main diagonal:

$$D = \begin{bmatrix} d_1 & 0 & 0 & \cdots & 0 \\ 0 & d_2 & 0 & \cdots & 0 \\ 0 & 0 & d_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & d_n \end{bmatrix},$$

where d_j, $j = 1, \ldots, n$ are some real numbers.

Special types of matrices: Identity matrix

The identity matrix is a diagonal matrix with every diagonal entry equal to 1:

$$I = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

It has the property that for any matrix $A \in \mathbb{R}^{n \times n}$, $IA = AI = A$.

Special types of matrices: Lower triangular matrix

A matrix $L \in \mathbb{R}^{m\times n}$ is lower triangular if all the entries above its main diagonal are zero. Square n-by-n lower triangular matrices take the form

$$L = \begin{bmatrix} \ell_{11} & 0 & 0 & \cdots & 0 \\ \ell_{21} & \ell_{22} & 0 & \cdots & 0 \\ \ell_{31} & \ell_{32} & \ell_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \ell_{n1} & \ell_{n2} & \ell_{n3} & \cdots & \ell_{nn} \end{bmatrix},$$

where $\ell_{i,j}$, $i = 1, \ldots, n$, $j = i, \ldots, n$ are some real numbers.

Special types of matrices: Upper triangular matrix

A matrix $U \in \mathbb{R}^{m\times n}$ is upper triangular if all the entries below its main diagonal are zero. Square n-by-n upper triangular matrices take the form

$$U = \begin{bmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1n} \\ 0 & u_{22} & u_{23} & \cdots & u_{2n} \\ 0 & 0 & u_{33} & \cdots & u_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & u_{nn} \end{bmatrix},$$

where $u_{i,j}$, $i = 1, \ldots, n$, $j = i, \ldots, n$ are some real numbers.

Special types of matrices: Symmetric matrix

A matrix A is symmetric if $A = A^T$. Note that only square matrices can be symmetric.

Inverse of a matrix

Let A be an n-by-n square matrix (i.e. $A \in \mathbb{R}^{n\times n}$). If there exists a square matrix $B \in \mathbb{R}^{n\times n}$ such that

$$BA = AB = I,$$

where I is the n-by-n identity matrix, then B is called the inverse of A.

- The inverse of A is denoted by A^{-1}.
- If A^{-1} exists then A is called nonsingular, otherwise it is singular.

Matrix inverses and linear systems

If A is a square, nonsingular matrix, then the solution to the linear system $A\mathbf{x} = \mathbf{b}$ is given formally as

$$\mathbf{x} = A^{-1}\mathbf{b}.$$

Important: When solving a linear system, one should never first compute A^{-1} and then compute the product $A^{-1}\mathbf{b}$**.** There are much better ways to solve the system (for example using Gaussian *elimination* when n is not too large)**.**

Properties of matrix inverses

Suppose A is nonsingular then the following statements are true

- A^{-1} is unique
- A^{-1} is nonsingular and its inverse is A
- A^T is nonsingular
- If $B \in \mathbb{R}^{n \times n}$ is nonsingular then AB is nonsingular and $(AB)^{-1} = B^{-1}A^{-1}$
- The linear system $A\mathbf{x} = \mathbf{b}$ has a unique solution.

Vector norms

- A vector norm is a scalar quantity that reflects the "size" of a vector $\mathbf{x}$.
- The norm of a vector $\mathbf{x}$ is denoted as $\|\mathbf{x}\|$.
- There are many ways to define the size of a vector. If $\mathbf{x} \in \mathbb{R}^n$, the three most popular are

$$\begin{aligned}
\text{one-norm}: &\quad \|\mathbf{x}\|_1 = \sum_{k=1}^{n} |x_k|, \\
\text{two-norm}: &\quad \|\mathbf{x}\|_2 = \sqrt{\sum_{k=1}^{n} |x_k|^2}, \\
\infty\text{-norm}: &\quad \|\mathbf{x}\|_\infty = \max_{1\le k\le n} |x_k|
\end{aligned}$$

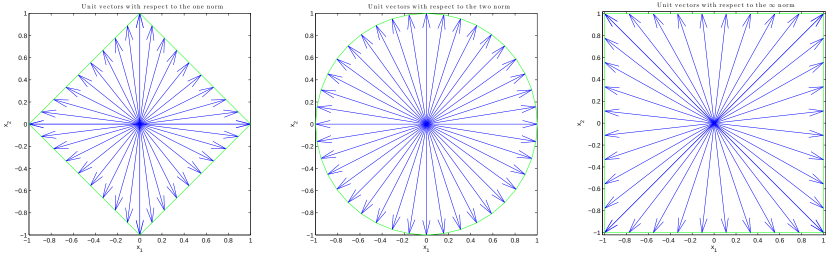
Vector norms

However, a vector norm is defined, it must satisfy the following three properties to be called a norm:

1. $\|\mathbf{x}\| \geq 0$ and $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = 0$ (i.e. $\mathbf{x}$ contains all zeros as its entries).
2. $\|\alpha\mathbf{x}\|$ = $|\alpha|\|\mathbf{x}\|$, for any constant α.
3. $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$, where $\mathbf{y} \in \mathbb{R}^n$. This is called the *triangle inequality*.

Unit vectors

- A vector $\mathbf{x}$ is called a unit vector if its norm is one, i.e. $\|\mathbf{x}\| = 1$.
- Unit vectors will be different depending on the norm applied.
- Below are several unit vectors in the one, two, and ∞ norms for $\mathbf{x} \in \mathbb{R}^2$.



(a) One-norm (b) Two-norm (c) ∞-norm

Matrix norms

- A matrix norm is a scalar quantity that reflects the "size" of a matrix $A \in \mathbb{R}^{m\times n}$.
- The norm of A is denoted as $\|A\|$.
- Any matrix norm must satisfy the following four properties:
  1. $\|A\| \geq 0$ and $\|A\| = 0$ if and only if $A = 0$ (i.e. A contains all zeros as its entries).
  2. $\|\alpha A\|$ = $|\alpha|\|A\|$, for any constant α.
  3. $\|A + B\| \leq \|A\| + \|B\|$, where $B \in \mathbb{R}^{m\times n}$.
  4. $\|AB\| \leq \|A\|\|B\|$, where $B \in \mathbb{R}^{n\times p}$. This is called the submultiplicative inequality.

Matrix norms

Each vector norm induces a matrix norm according to the following definition:

$$\|A\|_p = \max_{\|\mathbf{x}\|_p \neq 0} \frac{\|A\mathbf{x}\|_p}{\|\mathbf{x}\|_p} = \max_{\|\mathbf{x}\|_p = 1} \|A\mathbf{x}\|_p,$$

where $\mathbf{x} \in \mathbb{R}^n$ and $p = 1, 2, \ldots$.

Induced norms describe how the matrix stretches unit vectors with respect to that norm.

Induced matrix norms

Two popular and easy to define induced matrix norms are

$$\text{One-norm}: \quad \|A\|_1 = \max_{1\le k\le n} \sum_{j=1}^{m} |a_{jk}|,$$

$$\infty\text{-norm}: \quad \|A\|_\infty = \max_{1\le j\le m} \sum_{k=1}^{n} |a_{jk}|.$$

- The one-norm corresponds to the maximum of the one norm of every column.
- The ∞-norm corresponds to the maximum of the one norm of every row.

The two-norm of A is defined as the *largest eigenvalue* of the matrix $A^T A$. This is computationally expensive to compute.

Non-induced matrix norms

The most popular matrix norm that is not an induced norm is the *Frobenius* norm:

$$\|A\|_{\mathrm{F}} = \sqrt{\sum_{j=1}^{m}\sum_{k=1}^{n}|a_{jk}|^2}.$$

Important results on matrix norms

The following are some useful inequalities involving matrix norms. Here $A \in \mathbb{R}^{m\times n}$:

- $\|A\mathbf{x}\| \leq \|A\|\|\mathbf{x}\|$
- $\dfrac{1}{\sqrt{m}}\|A\|_1 \leq \|A\|_2 \leq \sqrt{n}\|A\|_1$
- $\dfrac{1}{\sqrt{n}}\|A\|_\infty \leq \|A\|_2 \leq \sqrt{m}\|A\|_\infty$
- $\|A\|_2 \leq \sqrt{\|A\|_1\|A\|_\infty}$