An Introduction to Linear Algebra: The Language of Chemometrics

Christian Damsgaard Jørgensen

Department of Mathematics and Computer Science
University of Southern Denmark

Week 38

Systems of Linear Equations

A *linear equation in n unknowns* is an equation of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

where $a_1, a_2, \ldots, a_n$ and b are in $\mathbb{R}$ and $x_1, x_2, \ldots, x_n$ are variables.

A *linear system* of m equations in n unknowns has the form

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & & & & & & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array} \tag{1}$$

where the a_{ij}'s and the b_i's are in $\mathbb{R}$.

We will refer to systems of the form (1) as $m \times n$ linear systems.

Systems of Linear Equations

A solution of an $m \times n$ system is an n-tuple of numbers

$$(x_1, x_2, \ldots, x_n)$$

that satisfies all the equations of the system.

An n-tuple can be interpreted as a vector of length n.

The set of all solutions of a linear system is called the *solution set*.

Consistency

- If the solution set is empty, the system is *inconsistent*.
- If the solution set is nonempty, the system is *consistent*.

To solve a consistent system, we must find its solution set.

Systems of Linear Equations

We will examine geometrically 2×2 systems of the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned}$$

Each equation can be represented graphically as a line in the plane.

An ordered pair (x_1, x_2) is a solution iff it lies on both lines.

In general, there are only three possible cases:

1. the lines are parallel,
2. the lines intersect at a point, or
3. both equations represent the same line

The solution set then contains either 0, 1, or infinitely many points.

Systems of Linear Equations

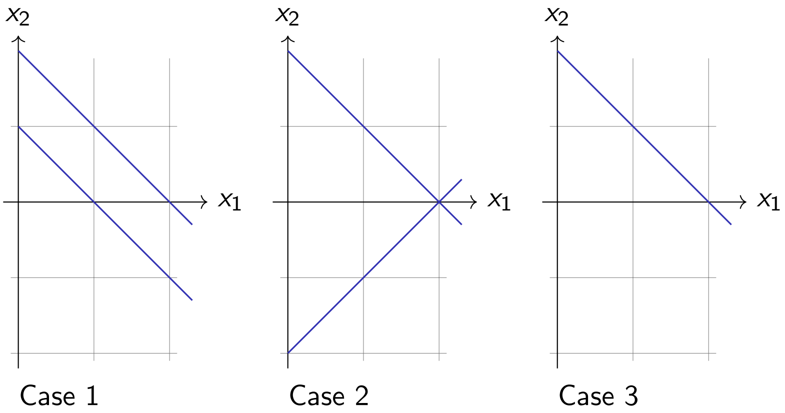

Figure 1: Geometric interpretation of solution sets.

Remark: The situation is the same for $m \times n$ systems.

Matrices and Vectors

Rectangular arrays are used to *store* information on computers. We will define algebraic operations on such arrays in a meaningful way.

Goal: Use arrays to *solve* linear systems of equations on computers.

A rectangular array of numbers with m rows and n columns as

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

is called an $m \times n$ matrix. We will often shorten this to $\mathbf{A} = \{a_{ij}\}$.

The number a_{ij} is called the (i, j) *entry* of the matrix $\mathbf{A}$.

Matrices and Vectors

We will use bold capital letters to refer to matrices.

Square Matrices

A matrix is *square* if it has the same number of rows and columns.

Triangular Matrices

A square matrix $\mathbf{A}$ is *upper triangular* if $a_{ij} = 0$ for $i > j$.

A square matrix $\mathbf{A}$ is *lower triangular* if $a_{ij} = 0$ for $i < j$.

Moreover, $\mathbf{A}$ is *triangular* if it is either upper or lower triangular.

Diagonal Matrices

A square matrix $\mathbf{A}$ is *diagonal* if $a_{ij} = 0$ for $i \neq j$.

Note: A diagonal matrix is both upper and lower triangular.

Matrices and Vectors

Diagonal Matrices (continued)

An entry a_{ij} of $\mathbf{A}$ is called a *main diagonal entry* if $i = j$.

The diagonal matrix with the main diagonal entries $d_1, d_2, \ldots, d_n$ is denoted by $\text{diag}\{d_1, d_2, \ldots, d_n\}$. A row vector consisting of the main diagonal entries of a square matrix $\mathbf{A}$ is denoted by $\text{diag}\{\mathbf{A}\}$.

Identity Matrices

An *identity matrix* is a square matrix $\mathbf{I} = \{\delta_{ij}\}$, where

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

Thus, it is a diagonal matrix whose main diagonal entries are 1.

The identity matrix is denoted by $\mathbf{I}$.

Matrices and Vectors

Zero Matrices

A *zero matrix* is a matrix whose entries are all 0. It is denoted by $\mathbf{0}$.

One Matrices

A *one matrix* is a matrix whose entries are all 1. It is denoted by $\mathbf{1}$.

$$\mathbf{I} = \underbrace{\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}}_{n \times n}, \mathbf{0} = \underbrace{\begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}}_{m \times n}, \mathbf{1} = \underbrace{\begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}}_{m \times n}$$

Figure 2: An identity matrix $\mathbf{I}$, a zero matrix $\mathbf{0}$, and a one matrix $\mathbf{1}$.

Matrices and Vectors

A vector turns out to be a special case of a matrix.

Vectors are one-dimensional arrays:

$$\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

These are matrices that consist of a single row or a single column.

Matrices that consist of a single row are called *row vectors*.

Matrices that consist of a single column are called *column vectors*.

We will denote the jth column vector of an $m \times n$ matrix $\mathbf{A}$ by $\mathbf{a}_j$.

Matrices and Vectors

Equality

Two $m \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$ are equal if $a_{ij} = b_{ij}$ for each i and j.

Scalar Multiplication

If $\mathbf{A}$ is an $m \times n$ matrix and α is a single number, then

$$\alpha\mathbf{A} = \alpha \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} \alpha a_{11} & \alpha a_{12} & \cdots & \alpha a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \cdots & \alpha a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha a_{m1} & \alpha a_{m2} & \cdots & \alpha a_{mn} \end{bmatrix}$$

is formed by multiplying each of the entries of $\mathbf{A}$ by α.

In this setting, the single number α is called a *scalar*.

Matrices and Vectors

The Transpose of a Matrix

The *transpose* of an $m \times n$ matrix $\mathbf{A} = \{a_{ij}\}$ is the $n \times m$ matrix $\mathbf{A}^\top$ whose (j, i) entry is a_{ij} for $j = 1, 2, \ldots, n$ and $i = 1, 2, \ldots, m$:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \mathbf{A}^\top = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}$$

The jth row of $\mathbf{A}^\top$ has the same entries as the jth column of $\mathbf{A}$.

The ith column of $\mathbf{A}^\top$ has the same entries as the ith row of $\mathbf{A}$.

Symmetric Matrices

A square matrix $\mathbf{A}$ is *symmetric* if $\mathbf{A}^\top = \mathbf{A}$.

Matrices and Vectors

Matrix Addition

If $\mathbf{A}$ and $\mathbf{B}$ are $m \times n$ matrices, then

$$\mathbf{A}+\mathbf{B} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}$$
$$= \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{bmatrix}$$

is formed by adding their corresponding entries.

Matrices and Vectors

Matrix Subtraction

If **A** and **B** are $m \times n$ matrices, then

$$\mathbf{A} - \mathbf{B} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} - \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} - b_{11} & a_{12} - b_{12} & \cdots & a_{1n} - b_{1n} \\ a_{21} - b_{21} & a_{22} - b_{22} & \cdots & a_{2n} - b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} - b_{m1} & a_{m2} - b_{m2} & \cdots & a_{mn} - b_{mn} \end{bmatrix}$$

is formed by subtracting the corresponding entry of **B** from each entry of **A**. Equivalently, we can define $\mathbf{A} - \mathbf{B}$ to be $\mathbf{A} + (-1)\mathbf{B}$.

Matrices and Vectors

Matrix Multiplication

If **A** is an $m \times n$ matrix and **B** is an $n \times r$ matrix, then

$$\mathbf{AB} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1r} \\ b_{21} & b_{22} & \cdots & b_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nr} \end{bmatrix}$$

$$= \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1r} \\ c_{21} & c_{22} & \cdots & c_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mr} \end{bmatrix} = \mathbf{C}$$

is an $m \times r$ matrix whose entries are defined by $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$.

In general, matrix multiplication is not commutative: $\mathbf{AB} \neq \mathbf{BA}$.

Matrices and Vectors

Linear Systems

We can represent an $m \times n$ system by a matrix equation of the form

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

$\mathbf{A}$ is an $m \times n$ matrix whose entries are variable coefficients, $\mathbf{x}$ is an $n \times 1$ vector of unknowns, and $\mathbf{b}$ is an $m \times 1$ right-hand side vector:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

The rule for determining the ith entry of the matrix product $\mathbf{A}\mathbf{x}$ is

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n$$

Matrices and Vectors

Linear Systems (continued)

The product **Ax** is defined by

$$\mathbf{Ax} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}$$

We can also express the product **Ax** as a sum of column vectors:

$$\mathbf{Ax} = x_1 \underbrace{\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}}_{\mathbf{a}_1} + x_2 \underbrace{\begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}}_{\mathbf{a}_2} + \cdots + x_n \underbrace{\begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}}_{\mathbf{a}_n}$$

Matrices and Vectors

Consistency Theorem for Linear Systems

An $m \times n$ system $\mathbf{Ax} = \mathbf{b}$ is consistent iff $\mathbf{b}$ can be written as

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n$$

which is called a *linear combination* of the column vectors of $\mathbf{A}$.

The Span of a Set of Vectors

The set of all linear combinations of the vectors is called their *span*.

Linear Independence

The vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ are linearly independent if

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{0}$$

implies that all the scalars $x_1, x_2, \ldots, x_n$ must equal 0.

Matrices and Vectors

Rank

The number of linearly independent rows or columns of an $m \times n$ matrix $\mathbf{A}$ is called the *rank* of $\mathbf{A}$ and is often denoted by $\text{rank}(\mathbf{A})$.

In general, $\text{rank}(\mathbf{A}) \leq \min(m, n)$.

The matrix $\mathbf{A}$ has *full rank*, if $\text{rank}(\mathbf{A}) = \min(m, n)$.

Singularity

A square matrix $\mathbf{A}$ is *nonsingular* or *invertible* if there exists a matrix $\mathbf{A}^{-1}$ (called the *multiplicative inverse* of $\mathbf{A}$), such that

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$$

If no such matrix exists, then $\mathbf{A}$ is singular.

We will often refer to $\mathbf{A}^{-1}$ as simply the *inverse* of $\mathbf{A}$.

Matrices and Vectors

Algebraic Rules

Each of the statements is valid for any scalars α and β and for any matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ for which the operations are defined:

1. $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
2. $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$
3. $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$
4. $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$
5. $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$
6. $(\alpha\beta)\mathbf{A} = \alpha(\beta\mathbf{A})$
7. $\alpha(\mathbf{AB}) = (\alpha\mathbf{A})\mathbf{B} = \mathbf{A}(\alpha\mathbf{B})$
8. $(\alpha + \beta)\mathbf{A} = \alpha\mathbf{A} + \beta\mathbf{A}$
9. $\alpha(\mathbf{A} + \mathbf{B}) = \alpha\mathbf{A} + \alpha\mathbf{B}$

Matrices and Vectors

Algebraic Rules for Transposes

There are four basic algebraic rules involving transposes:

1. $(\mathbf{A}^\top)^\top = \mathbf{A}$
2. $(\alpha \mathbf{A})^\top = \alpha \mathbf{A}^\top$
3. $(\mathbf{A} + \mathbf{B})^\top = \mathbf{A}^\top + \mathbf{B}^\top$
4. $(\mathbf{AB})^\top = \mathbf{B}^\top \mathbf{A}^\top$

The Usual Norm on $\mathbb{R}^n$

The *Euclidean norm* of an $n \times 1$ vector $\mathbf{x}$ is defined by

$$\|\mathbf{x}\| = \sqrt{\mathbf{x}^\top \mathbf{x}} = \sqrt{\sum_{i=1}^{n} x_i^2} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

Matrices and Vectors

The Usual Norm on $\mathbb{R}^n$ (continued)

Similarly, for an $1 \times n$ vector $\mathbf{x}$, it is defined by

$$\|\mathbf{x}\| = \sqrt{\mathbf{x}\mathbf{x}^\top} = \sqrt{\sum_{i=1}^{n} x_i^2} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

A vector $\mathbf{x}$ is a *unit vector* if its Euclidean norm $\|\mathbf{x}\|$ is 1.

A natural extension of the notion of a vector norm to matrices is

$$\|\mathbf{A}\|_F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2} = \sqrt{a_{11}^2 + a_{12}^2 + \cdots + a_{21}^2 + a_{22}^2 + \cdots}$$

where $\mathbf{A}$ is an $m \times n$ matrix. This is called the *Frobenius norm*.

Matrices and Vectors

Trace

The *trace* of an $n \times n$ matrix $\mathbf{A}$ is the sum of its diagonal entries:

$$\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii} = a_{11} + a_{22} + \cdots + a_{nn}$$

A matrix and its transpose have the same trace: $\mathrm{tr}(\mathbf{A}) = \mathrm{tr}(\mathbf{A}^\top)$.

Three useful properties of the trace are

- $\mathrm{tr}(\mathbf{A} + \mathbf{B}) = \mathrm{tr}(\mathbf{A}) + \mathrm{tr}(\mathbf{B})$
- $\mathrm{tr}(\alpha\mathbf{A}) = \alpha\,\mathrm{tr}(\mathbf{A})$
- $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$

where $\mathbf{A}$ and $\mathbf{B}$ are $n \times n$ matrices and α is a scalar.

Matrices and Vectors

Orthogonality in $\mathbb{R}^n$

Two $n \times 1$ vectors $\mathbf{x}$ and $\mathbf{y}$ are *orthogonal* if $\mathbf{x}^\top \mathbf{y} = 0$.

We will denote this relationship by $\mathbf{x} \perp \mathbf{y}$.

Orthogonal Sets

A set of vectors is orthogonal if the vectors are pairwise orthogonal.

Orthonormal Sets

An *orthonormal set* of vectors is an orthogonal set of unit vectors.

Orthogonal Matrices

An $n \times n$ matrix $\mathbf{A}$ whose column vectors form an orthonormal set is called *orthogonal*. Thus, the matrix $\mathbf{A}$ is orthogonal iff $\mathbf{A}^\top \mathbf{A} = \mathbf{I}$.

Matrices and Vectors

Properties of Orthogonal Matrices

The vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ form a basis for $\mathbb{R}^n$ iff

1. $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are linearly independent
2. every vector in $\mathbb{R}^n$ is a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$

If $\mathbf{A}$ is an $n \times n$ orthogonal matrix, then

1. the column vectors of $\mathbf{A}$ form an orthonormal basis for $\mathbb{R}^n$
2. $\mathbf{A}^\top \mathbf{A} = \mathbf{I}$
3. $\mathbf{A}^\top = \mathbf{A}^{-1}$
4. $(\mathbf{A}\mathbf{y})^\top (\mathbf{A}\mathbf{x}) = \mathbf{y}^\top \mathbf{A}^\top \mathbf{A}\mathbf{x} = \mathbf{y}^\top \mathbf{x}$
5. $\|\mathbf{A}\mathbf{x}\| = \|\mathbf{x}\|$

Literature

Steven J. Leon.
Linear Algebra with Applications (9th Edition).
Pearson, 2015.

Bent Jørgensen, Yuri Goegebeur.
Lecture Notes for *Module 2: Matrix Algebra*.
statmaster.sdu.dk, 2007.