A Short Review of Linear Algebra

Yu-Tzung Chang and Hsuan-Wei Lee

Department of Political Science, National Taiwan University

2018.09.27

Outline

- What is linear algebra? And why we need this field?
- Basic definitions, operations and properties
- Gauss elimination and matrix row operations
- Subspace and dimensionality
- Matrix functions
- Matrix decompositions (eigenvalue decomposition, SVD)

What is linear algebra?

- A math course in which we deal with a lot of *vectors* and *matrices*?
- A math course in which we learn to systematically solve *linear equations*?
- A math course in which we calculate *eigenvalues* and *determinants* of matrices?
- A math course in which we play with *vector spaces*, *inner products* and *linear transforms*?
- All of the above!

What is linear algebra?

Not logged in Talk Contributions Create account Log in

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

Interaction

Help
About Wikipedia
Community portal
Recent changes
Contact page

Tools

Article Talk Read Edit View history Search Wikipedia

Linear algebra

From Wikipedia, the free encyclopedia

Not to be confused with Elementary algebra.

Linear algebra is the branch of mathematics concerning linear equations such as

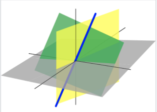
$$a_1x_1 + \cdots + a_nx_n = b,$$

linear functions such as

$$(x_1, \ldots, x_n) \mapsto a_1x_1 + \ldots + a_nx_n,$$

and their representations through matrices and vector spaces.[1][2][3]

Linear algebra is central to almost all areas of mathematics. For instance, linear algebra is fundamental in modern presentations of geometry, including for defining basic objects such as lines, planes and rotations. Also, functional analysis may be basically viewed as the application of linear algebra to spaces of functions. Linear algebra is also used in most sciences and engineering areas, because it allows modeling many natural phenomena, and efficiently computing with such models. For nonlinear systems, which cannot be modeled with linear algebra, linear algebra is often used as a first-order approximation.



In the three-dimensional Euclidean space, planes represent solutions of linear equations and their intersections represent the common solutions

- Linear algebra is the branch of mathematics concerning *vector spaces*, often finite or countably infinite dimensional, as well as *linear mappings* between such spaces.
- Such an investigation is initially motivated by a system of linear equations in several unknowns. Such equations are naturally represented using the formalism of *vectors* and *matrices*.

Why study linear algebra?

- Linear algebra is vital in multiple areas of science in general.
- First order approximation of the real and complex non-linear world.
- It's everywhere, e.g. physics, chemistry, biology, computer science, statistics, economics, social science, management, etc.

Vectors

- Vector $x \in \mathbb{R}^m$

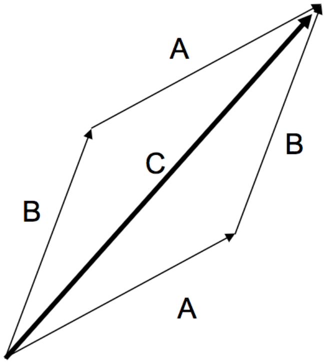
$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$

- May also write

$$x = \begin{bmatrix} x_1 & x_2 & \cdots & x_m \end{bmatrix}^T$$

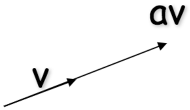
Vector addition

$$A + B = (x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$$



Scalar product

$$\alpha v = \alpha(x_1, x_2) = (\alpha x_1, \alpha x_2)$$



Matrices

- A matrix (plural: matrices) is a rectangular array of numbers, symbols, or expressions, arranged in rows and columns.
- Matrix $M \in \mathbb{R}^{m \times n}$

$$M = \begin{bmatrix} M_{11} & \dots & M_{1n} \\ \vdots & \ddots & \vdots \\ M_{m1} & \dots & M_{mn} \end{bmatrix}$$

- Written in terms of rows or columns

$$M = \begin{bmatrix} r_1^T \\ r_2^T \\ \vdots \\ r_m^T \end{bmatrix} = \begin{bmatrix} c_1 & c_2 & \dots & c_n \end{bmatrix}$$

where $r_i = \begin{bmatrix} M_{i1} & \dots & M_{in} \end{bmatrix}^T$, and $c_i = \begin{bmatrix} M_{1i} & \dots & M_{mi} \end{bmatrix}^T$

Multiplication

- Vector-vector: $x, y \in \mathbb{R}^m \to \mathbb{R}$

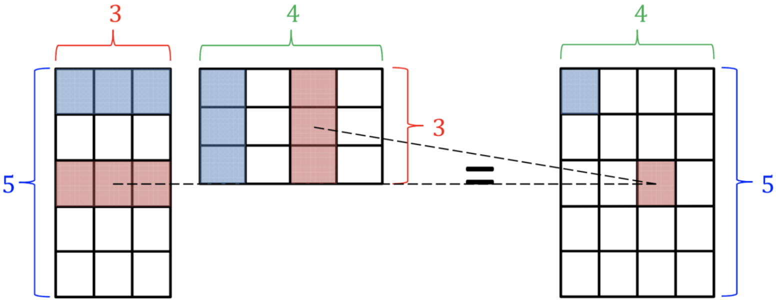
$$x^T y = \sum_{i=1}^{m} x_i y_i$$

- Matrix-vector: $x \in \mathbb{R}^n, M \in \mathbb{R}^{m \times n} \to \mathbb{R}^m$

$$Mx = \begin{bmatrix} r_1^T x \\ r_2^T x \\ \vdots \\ r_m^T x \end{bmatrix}$$

Multiplication (matrix-matrix)

- Matrix-matrix: $A \in \mathbb{R}^{m \times k}, B \in \mathbb{R}^{k \times n} \rightarrow \mathbb{R}^{m \times n}$



- Matrix-matrix: $A \in \mathbb{R}^{m \times k}, B \in \mathbb{R}^{k \times n} \rightarrow \mathbb{R}^{m \times n}$
- a_i: rows of A, b_j columns of B

$$AB = \begin{bmatrix} Ab_1 & \dots & Ab_n \end{bmatrix} = \begin{bmatrix} a_1^T B \\ \vdots \\ a_m^T B \end{bmatrix} = \begin{bmatrix} a_1^T b_1 & \dots & a_1^T b_n \\ \vdots & a_i^T b_j & \vdots \\ a_m^T b_1 & \dots & a_m^T b_n \end{bmatrix}$$

Multiplication properties

- *Associative*

$$(AB)C = A(BC)$$

- *Distributive*

$$A(B + C) = AB + AC$$

- NOT *commutative*

$$AB \neq BA$$

Gauss elimination and matrix operation

- Use Gauss elimination to solve the following linear system:

$$\begin{cases} x + y - z = 9 \\ y + 3z = 3 \\ -x - 2z = 2 \end{cases}$$

Useful matrices

- *Identity* matrix $I \in \mathbb{R}^{m \times m}$

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

$$AI = A, IA = A$$

- *Diagonal* matrix $A \in \mathbb{R}^{m \times m}$

$$A = diag(a_1, \dots, a_m) = \begin{bmatrix} a_1 & \dots & 0 \\ \vdots & a_i & \vdots \\ 0 & \dots & a_m \end{bmatrix}$$

Useful matrices

- *Symmetric* matrix: $A \in \mathbb{R}^{m \times m}$

$$A = A^T$$

- *Positive semidefinite* matrix $A \in \mathbb{R}^{m \times m}$:

$$x^T A x \geqslant 0, \quad \forall x \in \mathbb{R}^n$$

Equivalently, there exists $L \in \mathbb{R}^{m \times m}$ such that

$$A = LL^T$$

Orthogonal matrices

- For vectors $v, w \in \mathbb{R}^n$ (Euclidean space), the inner product of them is defined as:

$$v \cdot w = |v||w|cos\theta$$

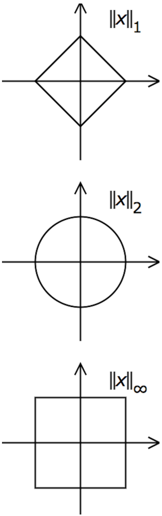
  where θ is the angle between the vectors.
- *Orthogonal* means perpendicular. Vectors v, w are orthogonal when the angle between them is $\theta = \frac{\pi}{2}$. Hence $cos\theta = 0$ and $v \cdot w = 0$.
- *Orthogonal* matrix: $U \in \mathbb{R}^{m \times m}$

$$U^T U = UU^T = I$$

Norms

- Quantify the "size" of a vector
- Given $x \in \mathbb{R}^n$, a norm satisfies
  - $||cx|| = |c|||x||$
  - $||x|| = 0 \iff x = 0$
  - $||x + y|| \leqslant ||x|| + ||y||$
- Common norms:
  - Euclidean L_2−norm: $||x||_2 = \sqrt{x_1^2 + \cdots + x_n^2}$
  - L_1−norm: $||x||_1 = |x_1| + \cdots + |x_n|$
  - L_∞−norm: $||x||_\infty = \max_i |x_i|$
  - L_p−norm: $\sqrt[p]{|x_1|^p + \cdots + |x_n|^p}$

norms illustration

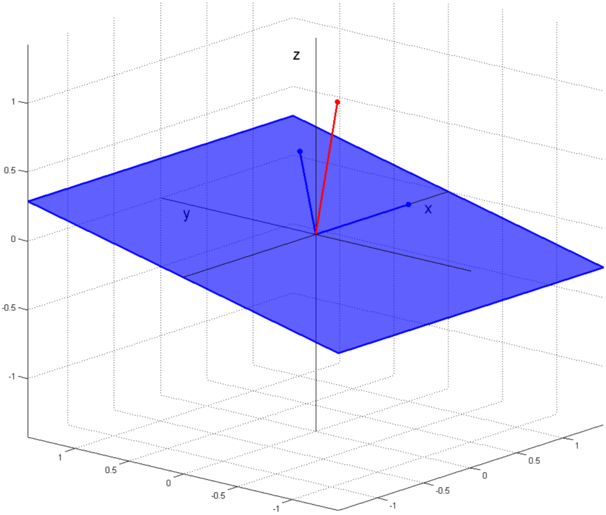


Linear subspace

Let K be a field (such as the real numbers), V be a vector space over K, and let W be a subset of V. Then W is a *subspace* if:

- The zero vector, 0, is in W.
- If u and v are elements of W, then the sum $u + v$ is an element of W.
- If u is an element of W and c is a scalar from K, then the scalar product cu is an element of W.

Linear subspace illustration



Linear span

The *span* of a set of S of vectors may be defined as the set of all finite linear combinations of elements of S.

$$span(S) = \left\{ \sum_{i=1}^{k} \lambda_i v_i \middle| k \in \mathbb{N}, v_i \in S, \lambda_i \in K \right\}$$

Linear independence

- The vectors $x_1, \dots, x_m$ in a subset S of a vector space V are said to be *linearly dependent*, if there exist scalars $a_1, a_2, \dots, a_k$, not all zero, such that

$$a_1 x_1 + a_2 x_2 + \cdots + a_k x_k = 0,$$

  where 0 denotes the zero vector.

- Vectors $x_1, \dots, x_m$ are *linearly independent* if

$$\sum_{i=1}^{m} \alpha_i x_i = 0 \iff \alpha_i = 0, \ \forall i$$

  Every linear combination of the x_i is unique.

Evaluating linear independence

Sets of vectors in R^2, check for linear dependency.

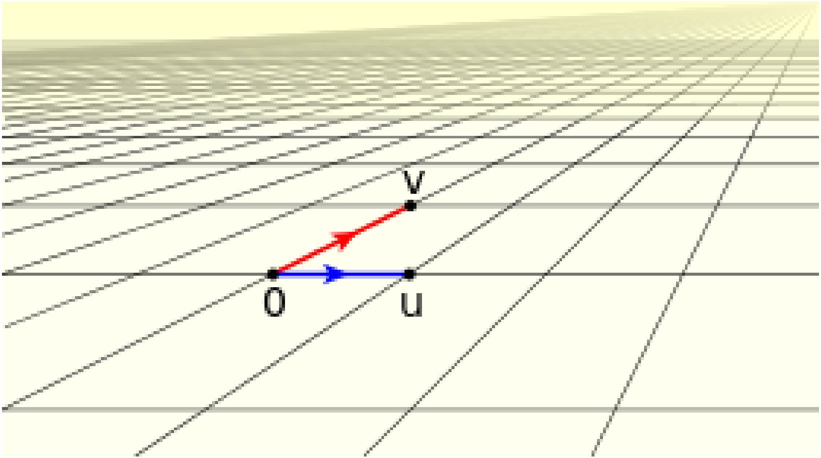
- The set of vectors $v_1 = (1, 1), v_2 = (-3, 2)$ and $v_3 = (2, 4)$.
- The set of vectors $v_1 = (1, 1), v_2 = (-3, 2)$.

Linear independence and dimension

- $Dim(V) = m$ if $x_1, \ldots, x_m$ span V and are linearly independent
- If $y_1, \ldots, y_k$ span V then
  - $k \geqslant m$
  - $k > m$ then y_i are NOT linearly independent

Linear Independence and Dimension

Dimension: Number of basis (linear independent) vectors of subspace for the subspace.

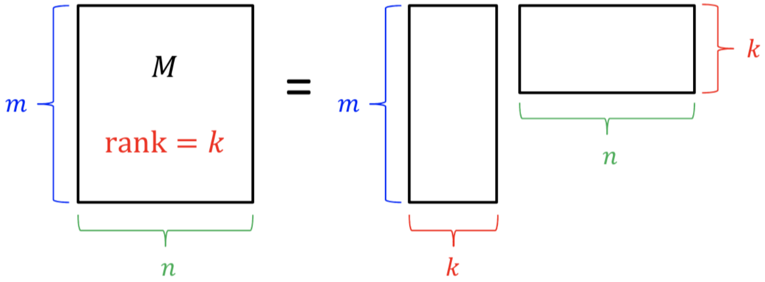


Matrix subspaces

- Matrix $M \in \mathbb{R}^{m \times n}$ defines two subspaces
  - Column space $col(M) = \{M\alpha \mid \alpha \in \mathbb{R}^n\} \subset \mathbb{R}^m$
  - Row space $row(M) = \{M^T\beta \mid \beta \in \mathbb{R}^m\} \subset \mathbb{R}^n$
- Nullspace of M: $null(M) = \{x \in \mathbb{R}^n \mid Mx = 0\}$
  - $null(M) \perp row(M)$
  - $dim(null(M)) + dim(row(M)) = n$
  - column space is similar

Matrix rank

- *rank*(M) gives *dimensionality* of row and column spaces
- If $M \in \mathbb{R}^{m \times n}$ has rank k, then it can be decomposed into product of $m \times k$ and $k \times n$ matrices



Example of matrix rank

- $$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$
- $$B = \begin{bmatrix} 1 & 5 & 3 & 5 & 9 & 1 \\ 2 & 6 & 4 & 4 & 8 & 2 \\ 3 & 7 & 5 & 3 & 7 & 4 \end{bmatrix}$$
- $$C = \begin{bmatrix} 3 & 3 & 3 \\ 3 & 3 & 3 \\ 3 & 3 & 3 \end{bmatrix}$$

Properties of rank

- For $A, B \in \mathbb{R}^{m \times n}$
  - $rank(A) \leqslant \min(m, n)$
  - $rank(A) = rank(A^T)$
  - $rank(AB) \leqslant \min(rank(A), rank(B))$, for any $B \in \mathbb{R}^{n \times k}$
  - $rank(A + B) \leqslant rank(A) + rank(B)$
- A has *full rank* if $rank(A) = \min(m, n)$
- If $m > rank(A)$ rows not linearly independent (similar for columns)

Matrix inverse

For a square matrix A, the inverse is written A^{-1}. When A is multiplied by A^{-1} the result is the identity matrix I. Non-square matrices do not have inverses.

Augmented matrix method to find the inverse of a matrix

$$\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right] \to \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{array}\right]$$

$$\to \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & 3/2 & -1/2 \end{array}\right]$$

$$\to \left[\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & 1 & 3/2 & -1/2 \end{array}\right]$$

Matrix inverse

- $M \in \mathbb{R}^{m \times m}$ is *invertible* iff $rank(M) = m$
- inverse is unique and satisfies
  - $M^{-1}M = MM^{-1} = I$
  - $(M^{-1})^{-1} = M$
  - $(M^T)^{-1} = (M^{-1})^T$
  - If A is invertible then MA is invertible and

$$(MA)^{-1} = A^{-1}M^{-1}$$

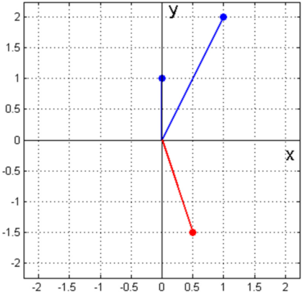
Systems of equations

- Given $M \in \mathbb{R}^{m \times n}$, $y \in \mathbb{R}^m$ we wish to solve

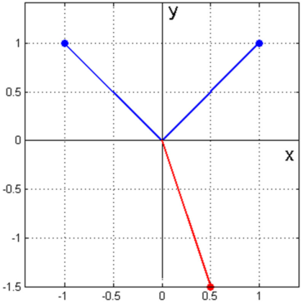
$$Mx = y$$

  The solution exists only if $y \in col(M)$
  (with possibly infinite number of solutions)
- If M is invertible then $x = M^{-1}y$
  Note: Do not solve M^{-1} by hands, use built-in functions...

Systems of equations (illustration)



$$\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} .5 \\ -2.5 \end{bmatrix} = \begin{bmatrix} .5 \\ -1.5 \end{bmatrix}$$



$$\begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ -.5 \end{bmatrix} = \begin{bmatrix} .5 \\ -1.5 \end{bmatrix}$$

Projection

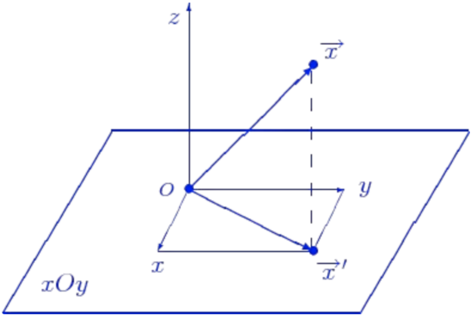
- What if $y \notin col(M)$?
- Find the x such that $\hat{y} = Mx$ is *closet* to y
  - $\hat{y}$ is the projection of y onto $col(M)$
  - also known as regression
- Assume $rank(M) = n < M$

$$x = (M^T M)^{-1} M^T y$$

$$\hat{y} = M(M^T M)^{-1} M^T y$$

Projection illustration

$M(M^T M)^{-1} M^T$ is a *projection matrix*



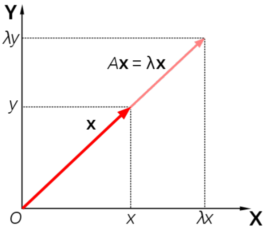
Eigenvalues and eigenvectors

An *eigenvector* or *characteristic vector* of a linear transformation is a non-zero vector that changes by only a scalar factor when that linear transformation is applied to it.

If T is a linear transformation from a vector space V over a field F into itself and v is a vector in V that is not the zero vector, then v is an *eigenvector* of T if $T(v)$ is a scalar multiple of v, that is,

$$T(v) = \lambda v.$$

And we call the scalar λ to be the *eigenvalue* corresponding to the eigenvector v.



Properties of eigenvalues

For $M \in \mathbb{R}^{m \times m}$ with eigenvalues λ_i

- $tr(M) = \sum_{i=1}^{m} \lambda_i$
- $det(M) = \lambda_1 \lambda_2 \ldots \lambda_m$
- $rank(M) = \# \lambda_i \neq 0$

When M is symmetric

- Eigenvalue decomposition is singular value decomposition
- Eigenvectors for nonzero eigenvalues give orthogonal basis for $row(M) = col(M)$

Eigenvalues decomposition

- Eigenvalues decomposition of symmetric $M \in \mathbb{R}^{m \times m}$ is

$$M = Q \Sigma Q^T = \sum_{i=1}^{m} \lambda_i q_i q_i^T$$

  - $\Sigma = diag(\lambda_1, \ldots, \lambda_n)$ contains eigenvalues of M
  - Q is orthogonal and contains eigenvectors q_i of M

Diagonalization

- Diagonalization problem: For a square matrix A does there exist an invertible matrix P such that $P^{-1}AP$ is diagonal?
- Diagonalizable matrix
  - Definition 1: A square matrix A is called *diagonalizable* if there exists an invertible matrix P such that $P^{-1}AP$ is a diagonal matrix (i.e., P diagonalizes A)
  - Definition 2: A square matrix A is called diagonalizable if A is similar to a diagonal matrix

Two square matrices A and B are *similar* if there exists an invertible matrix P such that $B = P^{-1}AP$.

Diagonalization example

Find A^k for

$$A = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix} = PDP^{-1}$$

We have

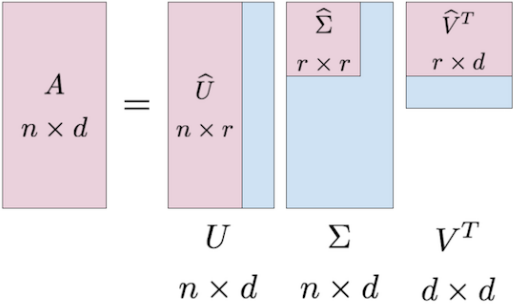
$$P = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}, \qquad D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$$

Then by associativity of matrix multiplication

$$A^2 = (PDP^{-1})(PDP^{-1}) = PD(P^{-1}P)DP^{-1} = PDDP^{-1} = PD^2P^{-1}$$

$$= \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$$

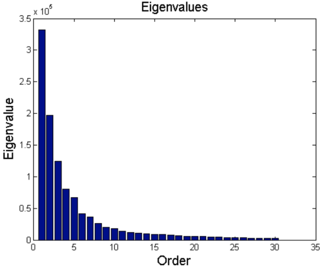
Single value decomposition (SVD)



- U: A real or complex orthonormal matrix
- V: A real or complex orthonormal matrix
- Σ : An $n \times d$ rectangular diagonal matrix with non-negative real numbers on the diagonal
- Interpretation: Use for dimensional reduction, find the "best" axis to project on. "Best": minimum sum of squares of projection errors.

SVD and image processing

Original picture



[1] https://www.projectrhea.org/rhea/index.php/PCA_Theory_Examples