Linear Algebra

Linear algebra is the branch of mathematics concerning vector spaces and linear mappings between such spaces.

1 Basic Definition

1.1 Matrix operation

- Addition: $(\mathbf{A}+\mathbf{B})_{i,j} = \mathbf{A}_{i,j} + \mathbf{B}_{i,j}$
- Scalar multiplication: $(c\mathbf{A})_{i,j} = c \cdot \mathbf{A}_{i,j}$
- Transposition: $(\mathbf{A}^{\mathrm{T}})_{i,j} = \mathbf{A}_{j,i}$
- Matrix multiplication: $(\mathbf{AB})_{i,j} = \sum_{r=1}^{n} A_{i,r}B_{r,j}$,
  - Associative: $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$
  - Distributive : $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$
  - Not commutative: $\mathbf{AB} \neq \mathbf{BA}$

1.2 Square Matrix

A **square matrix** is a matrix with the same number of rows and columns.

- Triangular: $\mathbf{U}_{i,j} = 0$ if $i > j$ and $\mathbf{L}_{i,j} = 0$ if $i < j$
- Diagonal (up. and low. tri.): $\mathbf{D}_{i,j} = 0$ if $i \neq j$
- Diagonalizable : $\mathbf{A} | \exists \mathbf{P}, \mathbf{P^{-1}AP} = \mathbf{D}$ (diagonal)
- Identity (diag.) : $\mathbf{I}_{i,j} = \{0 \text{ if } i \neq j \text{ ; } 1 \text{ if } i = j\}$
- Symmetric matrix: $\mathbf{A} = \mathbf{A}^{\mathrm{T}}$ or $a_{i,j} = a_{j,i}$
- Invertible: $\mathbf{A} \mid \exists \mathbf{A^{-1}}$ with $\mathbf{AA^{-1}} = \mathbf{A^{-1}A} = \mathbf{I}$
- Positive definite (sym.): $\mathbf{A} \mid \mathbf{x}^T\mathbf{Ax} > 0 \quad \forall \mathbf{x} \neq 0 \in \mathbb{R}^n$
- Orthogonal matrix: $\mathbf{A}_{i,.}$ and $hbfA_{.,j}$ are orthonormal vectot. This is equivalent to say $\mathbf{A}^{\mathrm{T}} = \mathbf{A^{-1}}$ or $\mathbf{A}^{\mathrm{T}}\mathbf{A} = \mathbf{I}$
- Unitary matrix (complex analogue to orthogonal): $\mathbf{A}^*\mathbf{A} = \mathbf{AA}^* = \mathbf{I}$

1.3 Main Operation

- Trace (square): $\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{n} \mathbf{A}_{i,i}$
- Determinant (square): is a number encoding certain properties of the matrix.
- The rank of a matrix is the maximum number of linearly independent columns (or line) vectors of the matrix.
- Conjugate: $(\overline{\mathbf{A}})_{i,j} = \mathrm{Re}(\mathbf{A}_{i,j}) - \mathrm{Im}(\mathbf{A}_{i,j})$
- Conjugate transpose or Hermitian transpose: $\boldsymbol{A}^* = \boldsymbol{A}^H = \overline{\boldsymbol{A}}^{\mathrm{T}} = \overline{\boldsymbol{A}^{\mathrm{T}}}$

2 Other stuff

2.1 Set, Field, Vector space, Span and Basis

- A *set* is a collection of distinct or well defined objects, considered as an object in its own right
- A *field* as a set together with two operations: addition and multiplication (axiom: associativity, commutativity and distributivity hold). Futhermore, every nonezero element of F has an inverse such that $f^{-1} \times f = 1$
- A *vector space* over a field F is a set V together with two operations: addition and scalar multiplication (the same axiom hold). The difference is that the multiplication is between an element of F and V and not two element of V. Any field F is a vector space over itself and in this special case the two operations are similar.
- The *kernel* (or nullspace) of a linear map $L : V \to W$, is the set of all elements v of V for which L(v) = 0. That is, in set-builder notation,

  $$\ker(L) = \{\mathbf{v} \in V | L(\mathbf{v}) = \mathbf{0}\}$$

  The difference of any two solutions u and v of the linear equation $\mathbf{Ax} = \mathbf{b}$ lies in the kernel of A $(\mathbf{u}-\mathbf{v}) \in \ker(A)$:

  $$\mathbf{A}(\mathbf{u}-\mathbf{v}) = \mathbf{Au} - \mathbf{Av} = \mathbf{b} - \mathbf{b} = \mathbf{0}$$

- The *linear span* of a set of vectors S in a vector space V is the intersection of all subspaces containing that set.
- A *basis* is a set of vectors (v_i)in a vector space V which are linearly independent (i.e. $\sum a_i v_i = 0 \Rightarrow a_i = 0$)and every other vector in the vector space is linearly dependent on these vectors (i.e. $x = \sum a_i v_i$).
- A *orthonormal basis* is a basis which vectors (e_i) are orthogonal and have unit length (i.e. inner product null $\langle e_i, e_j \rangle = 0$ if $i \neq j$ and $\langle e_i, e_i \rangle = \|e_i\| = 1$)
- The *support* of a function is the set of points where the function is not zero-valued :$\mathrm{supp}(f) = \{x \in X \mid f(x) \neq 0\}$
- The *frame* of a vector space V with an inner product (Hilbert space ?) can be seen as a generalization of the idea of a basis to sets which may be linearly dependent.

2.2 Linear dependence

A set of n vectors $\mathbf{v}_i$ is said to be linearly dependent if one of the vectors $\mathbf{v}_k$ in the set can be defined as a linear combination of the other vectors.

$$\mathbf{v}_k = \sum_{i=0}^{n} a_i \mathbf{v}_i \quad \text{with } i \neq k$$

And therefore the set of vector is linear independent if the only representations of 0 as a linear combination of its vectors is the trivial representation:

$$\sum_{i=0}^{n} a_i \mathbf{v}_i = 0 \Rightarrow \forall a_i = 0$$

- n vectors in $\mathbb{R}^n$ are linearly independent if and only if the determinant of the matrix formed by taking the vectors as its columns is non-zero.

$$\det([\mathbf{v}_1, ...\mathbf{v}_n]) = 0 \Leftrightarrow \mathbf{v}_1, ...\mathbf{v}_n \text{ are linearly independant}$$

2.3 Eigenvalues and eigenvectors

When viewing the square matrix A as a linear transformation, an **eigenvector** is a vector which direction is invariant under the transformation and which norm change by a scalar λ called eigenvalue.

$$\mathbf{Av} = \lambda \mathbf{v}$$

This can be rewrite $\mathbf{Av} - \lambda \mathbf{v} = (\mathbf{A} - \lambda \mathbf{I})\mathbf{v} = 0$. Either the solution is $x = 0$ in which case, the system would be inversible, be linearly indenpendant and has a none zero determinant. Either it is not the only solution, the system is singular and the determinant is equal to 0

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0$$

2.4 Inner Product

The **inner Product** associates a scalar quantity ($\in F$) to each pair of vectors ($\in V$): $\langle \cdot, \cdot \rangle : V \times V \to F$. It satisfies conjugate symmetry $\langle x, y \rangle = \overline{\langle y, x \rangle}$, linearity with first argument $\langle ax, y \rangle = a\langle x, y \rangle$ and positive-definiteness $\langle x, x \rangle \geq 0$.

- The inner product generalized the **dot product** in euclidean space: $\mathbf{A} \cdot \mathbf{B} = \|\mathbf{A}\| \, \|\mathbf{B}\| \cos\theta = \sum_{i=1}^{n} A_i B_i$
- **Orthogonality** is define by a inner product null
- **Norm** is define as the inner product of the vector itself

We say that two non-zero vectors u and v are **conjugate** (with respect to A) if and only if they are orthogonal with respect to this inner product:

$$\langle \mathbf{u}, \mathbf{v} \rangle_{\mathbf{A}} = \mathbf{u}^{\mathrm{T}} \mathbf{A} \mathbf{v} = 0$$

2.5 Complete Space

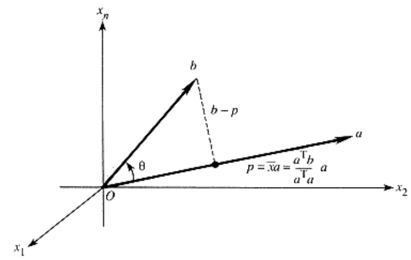
A metric space M is called complete if every Cauchy sequence (sequence whose elements become arbitrarily close to each other as the sequence progresses) of points in M has a limit that is also in M or, alternatively, if every Cauchy sequence in M converges in M.

2.6 Projection

In the general definition a projection is a linear transformation P from a vector space to itself such that $\mathrm{P}^2 = \mathrm{P}$ (idempotent). In a Hilbert space with the inner product, the orthogonal projection can be used

The projection of two vector can be written as

$$\mathrm{proj}_{\mathbf{u}}(\mathbf{v}) = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle} \mathbf{u} = \frac{\mathbf{u}\mathbf{u}^{\mathrm{T}}}{\mathbf{u}^{\mathrm{T}}\mathbf{u}} \mathbf{v} = \mathbf{Pv}$$



The projection $\mathbf{p}$ of a vector $\mathbf{b}$ into a subspace $\mathbf{A} = [\mathbf{a}_1, ...\mathbf{a}_n]$ must have the error $\mathbf{b} - \mathbf{p}$ perpendicular to the subspace

$$\mathbf{A}^{\mathrm{T}}(\mathbf{b} - \mathbf{p}) = 0 = \mathbf{A}^{\mathrm{T}}(\mathbf{b} - \mathbf{Ax}) \Rightarrow \mathbf{x} = (\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-\mathbf{1}}\mathbf{A}^{\mathrm{T}}\mathbf{b}$$

and,

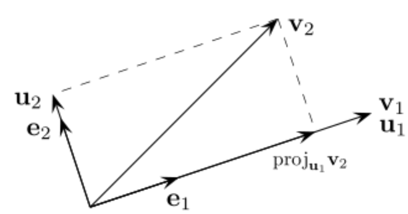
$$\mathbf{p} = \mathbf{Ax} = \mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-\mathbf{1}}\mathbf{A}^{\mathrm{T}}\mathbf{b}$$

Where $P = Ax = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$ is the projection matrix.

2.7 Gram-Schmidt Process

The Gram-Schmidt process is a method for orthonormalising a basis. It takes a finite, linearly independent set $v_1, \ldots, v_k$ and generates an orthonormal set $e_1, ..., e_k$.

$$\mathbf{u}_k = \mathbf{v}_k - \sum_{j=1}^{k-1} \mathrm{proj}_{\mathbf{u}_j}(\mathbf{v}_k) \qquad \mathbf{e}_k = \frac{\mathbf{u}_k}{\|\mathbf{u}_k\|}.$$



Numerical stability can be improve with the stabilized Gram-Schmidt

$i \leftarrow 1$
while $i < k$ **do**
  $\mathbf{v}_i \leftarrow \frac{\mathbf{v}_i}{\|\mathbf{v}_i\|}$ ▷ normalize
  $j \leftarrow i$
  for $j < k + 1$ **do**
    $\mathbf{v}_j \leftarrow \mathbf{v}_j - \mathrm{proj}_{\mathbf{v}_i}(\mathbf{v}_j)$ ▷ remove component in direction v_i
    $j \leftarrow j + 1$
  end for
  $i \leftarrow i + 1$
end while

2.8 Inversion

A square matrix $\mathbf{A}$ is **invertible** if it exist $\mathbf{A}^{-1}$ such that $\mathbf{AA^{-1}} = \mathbf{A^{-1}A} = \mathbf{I}$ It is equivalent to say:

- The columns of $\mathbf{A}$ are linearly independent.
- A has full rank; $\text{rank}(\mathbf{A}) = \dim(\mathbf{A})$.
- $\det(\mathbf{A}) \neq 0$
- Solution of $\mathbf{Ax} = \mathbf{b}$ is unique
- The only solution of $\mathbf{Ax} = \mathbf{0}$ is $\mathbf{x} = 0$

A not invertible matrix is called singular

2.9 Generalized and pseudo inverse

A **generalized inverse** of a matrix A is a matrix that has some properties of the inverse matrix of A but not necessarily all of them:

1. generalized inverse $AA^{\mathrm{g}}A = A$
2. A^g is a weak inverse $A^{\mathrm{g}}AA^{\mathrm{g}} = A^{\mathrm{g}}$
3. AA^g is Hermitian $(AA^{\mathrm{g}})^{\mathrm{T}} = AA^{\mathrm{g}}$
4. A^gA is Hermitian $(A^{\mathrm{g}}A)^{\mathrm{T}} = A^{\mathrm{g}}A$

If it satisfies all 4 conditions, then it is a (MoorePenrose) pseudoinverse. Some properties:

- The pseudoinverse exists and is unique for all matrix A.
- If A is invertible, its pseudoinverse is its inverse $A^+ = A^{-1}$.
- The pseudoinverse of the pseudoinverse is the original matrix: $(A^+)^+ = A$

A

2.10 Positive-Define Matrix

Any quadratic function can be written as $f(\mathbf{z}) = \mathbf{z}^{\mathrm{T}}\mathbf{Mz}$. The meaning of a positive define matrix, is one which the function associated has a unique minimum zero when $\mathbf{z} = 0$ and strictly positive otherwise and is strictly convex.

The following properties are equivalent to A, a symmetric square matrix being positive definite:

- $x^TAx > 0 \quad \forall x \neq 0 \in$
- All its eigenvalues are positive.
- all these determinants are positive

3 Matrix decomposition

A matrix decomposition is a factorization of a matrix into a product of matrices.

3.1 Decompositions related to solving systems of linear equations

3.1.1 LU factorization

of a square matrix $\mathbf{A}$ is into a lower L and upper U triangular matrix $\mathbf{A} = \mathbf{LU}$

3.1.2 QR factorization

of a matrix $\mathbf{A}$ with linear independent vector is into a product $\mathbf{A} = \mathbf{QR}$ of an orthogonal ($\mathbf{Q^TQ} = \mathbf{I}$) matrix $\mathbf{Q}$ and an upper triangular matrix $\mathbf{R}$ (or U). The orthonogal matrix can be produce with the Gram-schmidt algorithm (transforming linear independent vector in orthonormal one)

3.1.3 Cholesky decomposition

of a square, symmetric, positive definite matrix $\mathbf{A}$ into upper triangular with positive diagonal entries and its transporse $\mathbf{A} = \mathbf{U^TU}$

3.2 Decompositions based on eigenvalues and related concepts

3.2.1 Eigendecomposition

of a square matrix $\mathbf{A}$ with distinct eigenvectors into a square matrix $\mathbf{Q}$ whose column is the eigenvector of $\mathbf{A}$ and adiagonal matrix $\mathbf{\Lambda}$ whose diagonal elements are the corresponding eigenvalues.

$$\mathbf{A} = \mathbf{Q\Lambda Q}^{-1}$$

3.2.2 Takagi's factorization

of a square, complex, symmetric matrix A into a diagonal matrix and a unitary and its transpose.

$$A = VDV^T$$

This is the special case of the eigendecomposition where A is real-symmetric and therfore V is invertible $V^T = V^{-1}$. The geometrical interpretation is a rotation V^T, a scaling D and the back-rotation transform V

3.2.3 Singular Value Decomposition

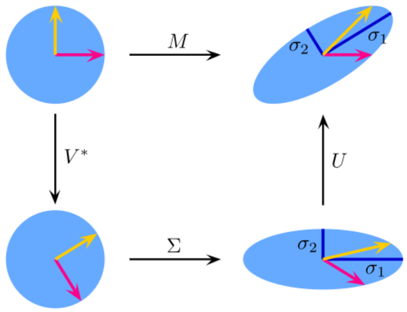
of an m n complex matrix M is a factorization of the form

$$\mathbf{M} = \mathbf{U\Sigma V}^*$$

where $\mathbf{U}$ is unitary, $\mathbf{\Sigma}$ is diagonal and $\mathbf{V}^*$ is the conjugate transpose of an unitary matrix $\mathbf{V}$. The diagonal entries σ_i of $\mathbf{\Sigma}$ are known as the singular values of $\mathbf{M}$.

Like the eigendecomposition below, SVD involves finding basis directions along which matrix multiplication is equivalent to scalar multiplication, but it has greater generality since the matrix under consideration need not be square.

In the case of square matrix with positive determinant, SVD can be interpreted as a composition of : a rotation ($\mathbf{V}^*$), a scaling ($\mathbf{\Sigma}$), and another rotation ($\mathbf{U}$).



$$M = U \cdot \Sigma \cdot V^*$$