MA60053 - Computational Linear Algebra
Lecture 2 - Revision of basics

M. Rajesh Kannan

Department of Mathematics,
Indian Institute of Technology Kharagpur,
email: rajeshkannan1.m@gmail.com, rajeshkannan@maths.iitkgp.ac.in

January 9, 2020

Notation

- $\mathbb{R}^n = \{(x_1, \ldots, x_n) : x_i \in \mathbb{R}\}$ and $\mathbb{C}^n = \{(x_1, \ldots, x_n) : x_i \in \mathbb{C}\}$.
- $\{e_1, \ldots, e_n\}$ - standard basis.
- For $x, y \in \mathbb{R}^n$, standard inner product $\langle x, y \rangle_2 = \sum_{i=1}^n x_i y_i$, and for $x, y \in \mathbb{C}^n$, standard inner product $\langle x, y \rangle_2 = \sum_{i=1}^n x_i \bar{y}_i$, $\bar{y}_i$ denotes the complex conjugate of y_i.
- $\mathbb{R}^{m \times n}$ denotes the set of all $m \times n$ matrices with real entries, and $\mathbb{C}^{m \times n}$ denotes the set of all $m \times n$ matrices with complex entries.

Inner product space

Definition

Let V be a vector space over $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$). An inner product is a function that assigns to every ordered pair of vector x and y in V, a scalar in $\mathbb{F}$, denoted by $\langle x, y\rangle$ such that $\forall$ x,$y \in V$, $\alpha \in \mathbb{F}$, the following hold

1. $\langle x + z, y\rangle = \langle x, y\rangle + \langle z, y\rangle$
2. $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$
3. $\langle x, y\rangle = \overline{\langle y, x\rangle}$
4. $\langle x, x\rangle > 0 \; \forall x \neq 0$

Inner product space

Definition

Let V be a vector space over $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$). An inner product is a function that assigns to every ordered pair of vector x and y in V, a scalar in $\mathbb{F}$, denoted by $\langle x, y \rangle$ such that $\forall$ x,y $\in$ V, $\alpha \in \mathbb{F}$, the following hold

1. $\langle x + z, y \rangle = \langle x, y \rangle + \langle z, y \rangle$
2. $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$
3. $\langle x, y \rangle = \overline{\langle y, x \rangle}$
4. $\langle x, x \rangle > 0 \; \forall x \neq 0$

Definition

A vector space V is an inner product space if there is an inner product defined on it.

Norm

Let V be a vector space over $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. A function $||.|| : V \times V \to [0, \infty)$ is a norm if for all $x, y \in V$ and $\lambda \in \mathbb{F}$, $||.||$ satisfies the following conditions:

- $||x|| = 0$ if and only if $x = 0$,
- $||\lambda x|| = |\lambda| ||x||$,
- $||x + y|| \leq ||x|| + ||y||$.

Example

- $V = \mathbb{R}$, $||x|| = |x|$, the absolute values of x.
- $V = \mathbb{R}^n$, $||x|| = \sqrt{\langle x, x\rangle}_2$.
- $V = \mathbb{R}^n$ and A be an $n \times n$ positive definite matrix, $||x||_A = \sqrt{\langle Ax, x\rangle}_2$ (Exercise)

Define $||x|| = \sqrt{\langle x, x \rangle}$

Theorem

Let V be an inner product space over $\mathbb{F}$. Then $\forall x, y \in V, c \in \mathbb{F}$, the following hold

1. $\|cx\| = |c|\|x\|$,
2. $\|x\| = 0$ *if and only if* $x = 0$, *and*
3. $\|x + y\| \leq \|x\| + \|y\|$.

Define $||x|| = \sqrt{\langle x, x \rangle}$

Theorem

Let V be an inner product space over $\mathbb{F}$. Then $\forall x, y \in V, c \in \mathbb{F}$, the following hold

1. $\|cx\| = |c|\|x\|$,
2. $\|x\| = 0$ *if and only if* $x = 0$, *and*
3. $\|x + y\| \leq \|x\| + \|y\|$.

Theorem (Cauchy-Schwarz Inequality)

$|\langle x, y \rangle| \leq ||x||||y|| \; \forall x, y \in V$

Definition

Let V be an inner product space. Two vectors x and y are said to be *orthogonal* if $\langle x, y \rangle = 0$.

Definition

Let V be an inner product space. Two vectors x and y are said to be *orthogonal* if $\langle x, y \rangle = 0$. A subset S of V is said to be *orthogonal* if any two distinct elements of S are orthogonal.

Definition

Let V be an inner product space. Two vectors x and y are said to be *orthogonal* if $\langle x, y\rangle = 0$. A subset S of V is said to be *orthogonal* if any two distinct elements of S are orthogonal. An orthogonal set S in V is said to be *orthonormal* if norm of every element in S is equal to 1.

Definition

Let V be an inner product space. Two vectors x and y are said to be *orthogonal* if $\langle x, y\rangle = 0$. A subset S of V is said to be *orthogonal* if any two distinct elements of S are orthogonal. An orthogonal set S in V is said to be *orthonormal* if norm of every element in S is equal to 1.

Definition (Orthonormal basis)

An ordered basis B of an inner product space V is said to be orthonormal basis if B is orthonormal.

Theorem

Let V be an inner product space and $S = \{v_1, v_2, \ldots, v_k\}$ be an orthogonal subset of V consisting of non zero vectors. If $y \in span(S)$, then
$y = \sum_{i=1}^{k} \frac{\langle y, v_i \rangle}{\|v_i\|^2} v_i$

Theorem

Let V be an inner product space and $S = \{v_1, v_2, \ldots, v_k\}$ be an orthogonal subset of V consisting of non zero vectors. If $y \in$ span(S), then $y = \sum_{i=1}^{k} \frac{\langle y, v_i \rangle}{\|v_i\|^2} v_i$

Remark

If $\{v_1, v_2, \ldots, v_n\}$ is an orthonormal set, then, for any $y \in$ span $\{v_1, v_2, \ldots, v_n\}$, we have $y = \sum_{i=1}^{k} \langle y, v_i \rangle v_i$

Theorem

Let V be an inner product space and $S = \{v_1, v_2, \ldots, v_k\}$ be an orthogonal subset of V consisting of non zero vectors. If $y \in span(S)$, then $y = \sum_{i=1}^{k} \frac{\langle y, v_i \rangle}{||v_i||^2} v_i$

Remark

If $\{v_1, v_2, \ldots, v_n\}$ is an orthonormal set, then, for any $y \in$ span $\{v_1, v_2, \ldots, v_n\}$, we have $y = \sum_{i=1}^{k} \langle y, v_i \rangle v_i$

Remark

If S is an orthonormal set in an inner product space V, consisting of non zero vectors, then S is linearly independent.

Definition (Gram-Schmidt Process)

Let V be an inner product space and $S = \{w_1, w_2, \ldots, w_n\}$ be a linearly independent set in V. Define $S^{'} = \{v_1, v_2, \ldots, v_n\}$, where $v_1 = w_1$ and $v_k = w_k - \sum_{i=1}^{k-1} \frac{\langle w_k, v_i \rangle . v_i}{||v_i||^2}, 2 \leq k \leq n$. Then $S^{'}$ is orthogonal and span(S) = span($S^{'}$)

Definition (Gram-Schmidt Process)

Let V be an inner product space and $S = \{w_1, w_2, \ldots, w_n\}$ be a linearly independent set in V. Define $S' = \{v_1, v_2, \ldots, v_n\}$, where $v_1 = w_1$ and $v_k = w_k - \sum_{i=1}^{k-1} \frac{\langle w_k, v_i \rangle . v_i}{\|v_i\|^2}, 2 \leq k \leq n$. Then S' is orthogonal and span(S) = span(S')

Theorem (Projection Theorem)

Let W be a finite dimensional subspace of an inner product space V and let $y \in V$. Then $\exists$ unique vector $u \in W$, $z \in W^{\perp}$ such that $y = u + z$. Furthermore if $\{v_1, v_2, \ldots, v_n\}$ is an orthonormal basis for W, then $u = \sum_{i=1}^{n} \langle y, v_i \rangle . v_i$

Definition (Gram-Schmidt Process)

Let V be an inner product space and $S = \{w_1, w_2, \ldots, w_n\}$ be a linearly independent set in V. Define $S' = \{v_1, v_2, \ldots, v_n\}$, where $v_1 = w_1$ and $v_k = w_k - \sum_{i=1}^{k-1} \frac{\langle w_k, v_i \rangle . v_i}{\|v_i\|^2}, 2 \leq k \leq n$. Then S' is orthogonal and span(S) = span(S')

Theorem (Projection Theorem)

Let W be a finite dimensional subspace of an inner product space V and let $y \in V$. Then $\exists$ unique vector $u \in W, z \in W^{\perp}$ such that $y = u + z$. Furthermore if $\{v_1, v_2, \ldots, v_n\}$ is an orthonormal basis for W, then $u = \sum_{i=1}^{n} \langle y, v_i \rangle . v_i$

Theorem (Riesz Representation Theorem)

Let V be a finite dimensional inner product space and let $g : V \to \mathbb{F}$ is linear. Then $\exists$ a unique vector $z \in V$ such that $g(x) = \langle x, z \rangle$

Four fundamental subspaces

For an $m \times n$ matrix A, the following subspaces are called fundamental subspaces.

- ***Range space of** A**:*** $R(A) = \{x \in \mathbb{R}^m : x = Ay \text{ for some } y \in \mathbb{R}^n\}$. (span of columns of A)

Four fundamental subspaces

For an $m \times n$ matrix A, the following subspaces are called fundamental subspaces.

- ***Range space of** A**:*** $R(A) = \{x \in \mathbb{R}^m : x = Ay \text{ for some } y \in \mathbb{R}^n\}$. (span of columns of A)
- ***Null space of** A**:*** $N(A) = \{x \in \mathbb{R}^n : Ax = 0\}$.

Four fundamental subspaces

For an $m \times n$ matrix A, the following subspaces are called fundamental subspaces.

- ***Range space of A:*** $R(A) = \{x \in \mathbb{R}^m : x = Ay \text{ for some } y \in \mathbb{R}^n\}$. (span of columns of A)
- ***Null space of A:*** $N(A) = \{x \in \mathbb{R}^n : Ax = 0\}$.
- ***Range space of A^T:*** $R(A^T) = \{x \in \mathbb{R}^n : x = A^T y \text{ for some } y \in \mathbb{R}^m\}$.

Four fundamental subspaces

For an $m \times n$ matrix A, the following subspaces are called fundamental subspaces.

- ***Range space of*** A***:*** $R(A) = \{x \in \mathbb{R}^m : x = Ay$ for some $y \in \mathbb{R}^n\}$. (span of columns of A)
- ***Null space of*** A***:*** $N(A) = \{x \in \mathbb{R}^n : Ax = 0\}$.
- ***Range space of*** A^T***:*** $R(A^T) = \{x \in \mathbb{R}^n : x = A^T y$ for some $y \in \mathbb{R}^m\}$.
- ***Null space of*** A^T***:*** $N(A^T) = \{x \in \mathbb{R}^m : A^T x = 0\}$.

Symmetric and Hermitian matrices

Definition

An $n \times n$ real matrix is said to be symmetric, if $A^T = A$.

Symmetric and Hermitian matrices

Definition

An $n \times n$ real matrix is said to be symmetric, if $A^T = A$. An $n \times n$ matrix is said to be Hermitian, if $A^* = A$.

Symmetric and Hermitian matrices

Definition

An $n \times n$ real matrix is said to be symmetric, if $A^T = A$. An $n \times n$ matrix is said to be Hermitian, if $A^* = A$. An $n \times n$ matrix is said to be normal, if $AA^* = A^*A$.

Symmetric and Hermitian matrices

Definition

An $n \times n$ real matrix is said to be symmetric, if $A^T = A$. An $n \times n$ matrix is said to be Hermitian, if $A^* = A$. An $n \times n$ matrix is said to be normal, if $AA^* = A^*A$.

Theorem

Eigenvalues of any Hermitian matrix are real numbers. (Eigenvalues of any real symmetric matrix are real numbers)

Symmetric and Hermitian matrices

Definition

An $n \times n$ real matrix is said to be symmetric, if $A^T = A$. An $n \times n$ matrix is said to be Hermitian, if $A^* = A$. An $n \times n$ matrix is said to be normal, if $AA^* = A^*A$.

Theorem

Eigenvalues of any Hermitian matrix are real numbers. (Eigenvalues of any real symmetric matrix are real numbers)

A real matrix A is said to be orthogonal if $AA^T = A^TA = I$, and a complex matrix A is said to be unitary if $AA^* = A^*A = I$.

Theorem (Schur, Jacobi)

Every complex matrix A is unitarily similar to an upper triangular matrix.

Theorem (Schur, Jacobi)

Every complex matrix A is unitarily similar to an upper triangular matrix.

Theorem (Spectral theorem for real symmetric matrices)

Any real symmetric matrix is orthogonally similar to a diagonal matrix.

Theorem (Schur, Jacobi)

Every complex matrix A is unitarily similar to an upper triangular matrix.

Theorem (Spectral theorem for real symmetric matrices)

Any real symmetric matrix is orthogonally similar to a diagonal matrix.

Theorem (Spectral theorem for Hermitian matrices)

Any Hermitian matrix is unitarily similar to a real diagonal matrix.

Theorem (Schur, Jacobi)

Every complex matrix A is unitarily similar to an upper triangular matrix.

Theorem (Spectral theorem for real symmetric matrices)

Any real symmetric matrix is orthogonally similar to a diagonal matrix.

Theorem (Spectral theorem for Hermitian matrices)

Any Hermitian matrix is unitarily similar to a real diagonal matrix.

Theorem (Spectral theorem for Normal matrices)

An $n \times n$ matrix A is normal if and only if A is unitarily similar to a diagonal matrix.

Spectral decomposition

Theorem

Let A be an $n \times n$ Hermitian matrix with rank r. Then A can be represented in each of the following equivalent forms:

- *There exists a unitary matrix P and a real diagonal nonsingular matrix Δ of rank r such that $A = P \begin{pmatrix} \Delta & 0 \\ 0 & 0 \end{pmatrix} P^*$.*

Spectral decomposition

Theorem

Let A be an $n \times n$ Hermitian matrix with rank r. Then A can be represented in each of the following equivalent forms:

- *There exists a unitary matrix P and a real diagonal nonsingular matrix Δ of rank r such that $A = P \begin{pmatrix} \Delta & 0 \\ 0 & 0 \end{pmatrix} P^*$.*
- *There exists non-zero real numbers $\lambda_1, \lambda_2, \ldots, \lambda_r$ and orthogonal vectors $u_1, \ldots, u_r$ such that $A = \sum_{i=1}^{n} \lambda_i u_i u_i^*$.*

Spectral decomposition

Theorem

Let A be an $n \times n$ Hermitian matrix with rank r. Then A can be represented in each of the following equivalent forms:

- *There exists a unitary matrix P and a real diagonal nonsingular matrix Δ of rank r such that* $A = P \begin{pmatrix} \Delta & 0 \\ 0 & 0 \end{pmatrix} P^*$.
- *There exists non-zero real numbers $\lambda_1, \lambda_2, \ldots, \lambda_r$ and orthogonal vectors $u_1, \ldots, u_r$ such that* $A = \sum_{i=1}^{n} \lambda_i u_i u_i^*$.
- *There exists matrices R and Δ of orders $n \times r$ and $r \times r$, respectively, such that Δ is real, diagonal and non-singular,* $R^* R = I$ *and* $A = R \Delta R^*$.

Positive Semidefinite Matrices(PSD)

Let $\mathcal{S}^n$ denote the subspace of symmetric matrices in $\mathbb{R}^{n\times n}$. $A \in \mathcal{S}^n$ is *positive semidefinite*(PSD) if $x^T Ax \geq 0$ for every $x \in \mathbb{R}^n$.

Theorem

TFAE for $A \in \mathcal{S}^n$:

(a) *A is PSD*

(b) *All the eigenvalues of A are nonnegative,*

(c) *All the principal minors of A are nonnegative,*

(d) *There exists an $n \times k$ real matrix B such that $A = BB^T$,*

(e) *There exists $C \in \mathcal{S}^n$ such that $A = C^2$,*

(f) *There exists an $n \times n$ lower triangular matrix L such that $A = LL^T$,*

(g) *There exists a k-dimensional Euclidean vector space V and vectors $v_1, \ldots v_n \in V$ such that $a_{ij} =< v_i, v_j >$,*

(h) *There exists k vectors $b_1, \ldots, b_k \in \mathbb{R}^n$ such that $A = \sum_{i=1}^{k} b_i b_i^k$.*

Positive definite matrices

$A \in \mathcal{S}^n$ is called *positive definite* (pd), if $x^T Ax > 0$ for every non zero $x \in \mathbb{R}^n$.

Theorem

TFAE for $A \in \mathcal{S}^n$:

(a) *A is pd,*

(b) *All the eigenvalues of A are positive,*

(c) *All the principal minors of A are positive,*

(d) *$A = BB^T$ for some nonsingular matrix B,*

(e) *$A = LL^T$, where L is a nonsingular lower triangular matrix,*

(f) *$A = C^2$ where $C \in \mathcal{S}^n$ is nonsingular,*

(g) *A is the Gram matrix of n linearly independent vectors,*

(h) *$A = \sum_{i=1}^{n} b_i b_i^T$, where $b_1, \ldots, b_n \in \mathbb{R}^n$ are linearly independent,*

(i) *A has a set of n positive nested principal minors, for example*

$$a_{11} > 0, \quad det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} > 0, \quad det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} > 0, \ldots, det(A) > 0$$

Schur complement

Definition

Let $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ *be a block matrix such that* D *is invertible. The Schur complement of* D *in* M *is, denote by* (M/D)*, defined by*

$$A - BD^{-1}C.$$

Schur complement

Definition

Let $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ *be a block matrix such that* D *is invertible. The Schur complement of* D *in* M *is, denote by* (M/D)*, defined by*

$$A - BD^{-1}C.$$

Motivation: Gaussian elimination for $\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} c \\ d \end{pmatrix}$.[With (M/D) invertible]

Properties

- $\det M = \det(M/D)\det(D)$,
- $\operatorname{rank} M = \operatorname{rank}(M/D) + \operatorname{rank} D$,
- Let M be symmetric, and D is nonsingular. Then, $M = \begin{pmatrix} A & B \\ B^T & D \end{pmatrix}$ is p.d. if and only if (M/D) and D are p.d.
- Let M be symmetric, and D is nonsingular. If D is p.d., then M is psd if and only if (M/D) are psd.