LA2 Summary

KYB

Thrn, it's a Fact

mathrnfact@gmail.com

December 4, 2020

Overview

Ajoint

X,U finite dimensional inner product space over $\mathbb{R}$ or $\mathbb{C}$ with basis $\mathcal{X} = \{x_1, \cdots, x_n\}$, $\mathcal{U} = \{u_1, \cdots, u_m\}$. For every linear map $T : X \to U$, there is a unique linear map $T^* : U \to x$ such that

$$\langle T(x), u \rangle_U = \langle x, T^*(u) \rangle_U \text{ for all } x \in X, u \in U.$$

How to compute

1. Let $M_{ij} = \langle u_j, T(x_i) \rangle_U$ and $G_{ij} = \langle x_j, x_i \rangle_X$.
2. $B = G^{-1}M$.
3. $[T^*] = B$.

Projection

H is a Hilbert space, S is a closed subspace oh H. Let $v \in H$.

- Then there is a unique vector $w \in S$ such that

$$\|v - w\|_2 = \min\{\|v - z\|_2 : z \in S\}.$$

  Denote $w = \operatorname{proj}_S v$.
- $w = \operatorname{proj}_S v$ iff $\langle v - w, z\rangle = 0$ for all $z \in S$.
- If S is finite dimensional, let $\{u_1, \cdots, u_n\}$ be a basis for S.
  1. $G_{ij} = \langle u_j, u_i\rangle$ and $b_i = \langle v, u_i\rangle$.
  2. Let $(x_1, \cdots, x_n) = G^{-1}b$. Then $\operatorname{proj}_S v = \sum_{i=1}^{n} x_i u_i$.

Least square solution

Let $A \in \mathbb{R}^{m\times n}$(resp. $\mathbb{C}^{m\times n}$) and $y \in \mathbb{R}^m$(resp. $\mathbb{C}^m$).

- Then there is a least square solution x, that is,

$$\|Ax - y\|_2 = \min\{\|Az - y\|_2 : z \in \mathbb{R}^n(\text{resp. } \mathbb{C}^n\}.$$

- x is a least square solution iff x satisfies $A^T Ax = A^T y$(resp.$A^* Ax = A^* y$).

Minimum norm least square solution

A least square solution $\overline{x}$ to $Ax = y$ is called the minimum norm least square solution if

$$\|\overline{x}\|_2 = \min\{\|x\|_2 : x \text{ is a least square solution to } Ax = y\}.$$

- The MNLS is unique.
- $\overline{x}$ is MNLS iff $A^T A\overline{x} = A^T y$ and $\overline{x} \in \mathrm{col}(A^T)$(resp.$A^* A\overline{x} = A^* y$ and $\overline{x} \in \mathrm{col}(A^*)$).

Orthogonal basis

Let $\{u_1, \cdots, u_n\}$ be an orthogonal basis of X and let $x \in X$. Then

$$x = \sum_{i=1}^{n} \frac{\langle x, u_i \rangle}{\langle u_i, u_i \rangle} u_i.$$

If $\{u_1, \cdots, u_n\}$ is orthonormal,

$$x = \sum_{i=1}^{n} \langle x, u_i \rangle u_i.$$

Gram-Schmidt

Given linearly independent set $\{u_1, \cdots, u_n\}$, $\{\hat{u}_1, \cdots, \hat{u}_n\}$ is orthogonal where

$$\hat{u}_1 = u_1,$$

$$\hat{u}_{k+1} = u_{k+1} - \sum_{i=1}^{k} \frac{\langle u_{k+1}, u_i \rangle}{\langle u_i, u_i \rangle} u_i.$$

If X is a finite inner product space, then there is an orthonormal basis.

Projection using orthonormal set

Let S be a finite dimensional subspace of an inner product space X. Let $\{u_1, \cdots, u_n\}$ be an orthonormal basis. For $v \in X$,

$$\operatorname{proj}_S v = \sum_{i=1}^{n} \langle v, u_i \rangle u_i.$$

Orthogonal complements

Let H be a Hilbert space and let S be a nonempty subset of H.

- $S^{\perp} = \{v \in V : \langle v, s \rangle = 0 \text{ for all } s \in S\}$ is a subspace of V.
- If S is a closed subspace, then $S^{\perp\perp} = S$.

Fundamental Theory of Linear Algebra

Let X,U finite dimensional inner product spaces and let $T : X \to U$ be linear.

- $\ker(T)^{\perp} = \mathcal{R}(T^*)$ and $\mathcal{R}(T^*)^{\perp} = \ker(T)$
- $\ker(T^*)^{\perp} = \mathcal{R}(T)$ and $\mathcal{R}(T)^{\perp} = \ker(T^*)$
- $\operatorname{rank}(T) = \operatorname{rank}(T^*)$
- $X = \mathcal{R}(T^*) \oplus \ker(T)$.

Spectral Decomposition

Let a matrix $A \in \mathbb{C}^{n \times n}$ be normal ($A^*A = AA^*$). Then A has a spectral decomposition, i.e. there are n distinct eigenpairs (λ_i, x_i) such that $\{x_1, \cdots, x_n\}$ is an orthonormal basis of $\mathbb{C}^n$ and

$$A = XDX^*$$

where $X = [x_1|\cdots|x_n]$ and $D = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$.

Hermitian matrices

Let $A \in \mathbb{C}^{n\times n}$ be Hermitian ($A^* = A$).

- Then every eigenvalue of A is real and $\mathrm{m.\,geo}(\lambda) = \mathrm{m.\,alg}(\lambda)$ for all eigenvalue λ.
- If $A \in \mathbb{R}^{n\times n}$, for each eigenvalue, there is a corresponding eigenvector x in $\mathbb{R}^n$.
- A is positive definite iff every eigenvalue is positive.

Optimization

Let $q : \mathbb{R}^n \to \mathbb{R}$ of the form

$$q(x) = \frac{1}{2}x \cdot Ax + b \cdot x + c.$$

We can always assume A is symmetric ($A_{\text{sym}} = \frac{1}{2}(A + A^T)$).

- If A is not positive semidefinite, $q(x)$ has no minimizer.
- If A is positive definite, $q(x)$ has a unique minimizer.
- Suppose A is positive semidefinite but not positive definite.
  - if $b \in \operatorname{col}(A)$, then $q(x)$ has a minimizer.
  - if $b \notin \operatorname{col}(a)$, then $q(x)$ has no minimizer.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a smooth function. If $x^* \in \mathbb{R}^n$

- $\nabla f(x^*) = 0$ and
- $\nabla^2 f(x^*)$ is positive semidefinite,

then x^* is a (local) minimizer of $f(x)$.

The SVD

Let $A \in \mathbb{C}^{m \times n}$. An eigenvalue of A^*A is called a singular value of A. Let $\sigma_1 \geq \cdots \geq \sigma_{\min\{m,n\}}$. Then there exists unitary matrices $U \in \mathbb{C}^{m \times m}$ and $V \in \mathbb{C}^{n \times n}$ and a diagonal matrix $\Sigma \in \mathbb{C}^{m \times n}$ with $\Sigma_{ii} = \sigma_i$ such that

$$A = U \Sigma V^*.$$

Let r be a positive integer such that $\sigma_r > 0$ but $\sigma_{r+1} = 0$. Then $\mathrm{rank}(A) = r$ and there is a reduced SVD

$$A = U_1 \Sigma_1 V_1^*.$$

Using outer product form

$$A = \sum_{i=1}^{r} \sigma_i u_i \otimes v_i.$$

LU factorization

Let $A \in \mathbb{R}^{n \times n}$.

- If there is no pivoting when applying Gaussian elimination to A, A has an LU factorization where L is a unit lower triangular matrix and U is an upper triangular matrix.
- If every submatrix $M^{(k)} \in \mathbb{R}^{k \times k}$ from the upper left-hand corner of A for $k = 1, \cdots, n-1$ is nonsingular, LU factorization is unique.

The Cholesky factorization

Let $A \in \mathbb{R}^{n \times n}$ be symmetric and positive definite.

- A has a LU factorization.
- the diagonal entries of U are all positive.
- $A = LDL^T = R^T R$.

Partial Pivoting

When applying Gaussian elimination to A, we may permute rows so that the leading coefficient is nozero. So

$$U = L_{n-1}P_{n-1}L_{n-2}P_{n-2}\cdots L_1P_1A,$$

where L_i are unit lower triangular matrices and P_i are either I or permuting $e_i \leftrightarrow e_j$ for some $j > i$. Let $P = P_{n-1}\cdots P_1$ and define

$$\tilde{L}_k = P_{n-1}P_{n-2}\cdots P_{k+1}L_kP_{k+1}P_{k+2}\cdots P_{n-1}.$$

Then

- $\tilde{L}_i$ is a unit lower triangular matrix.
- $\tilde{L}_{n-1}\tilde{L}_{n-2}\cdots\tilde{L}_1PA = U$.

Proof

Note that $P_i^2 = I$ for all i. Thus the second result is trivial.
Each L_i is a unit lower triangular matrix such that

$$L_i e_k = e_k \text{ if } k \neq i.$$

Each P_i is a permutation of rows such that

$$P_i e_i = e_j, P_i e_j = e_i, P_i e_k = e_k \ (k \neq i, j)$$

for some $j \geq i$ (if $i = j$, $P_i = I$). Thus it suffices to show that

- $\tilde{L}_i e_k = e_k$ for all $k \neq i$,
- $\tilde{L}_i e_i = (0, \cdots, 0, 1, \cdots)$.

(continued)

Proof

Let $k < i$. Then

$$\tilde{L}_i e_k = P_{n-1}P_{n-2}\cdots P_{i+1}L_kP_{i+1}P_{i+2}\cdots P_{n-1}e_k = e_k.$$

Let $k > i$. Put $k_1 = k$ and $e_{k_{j+1}} = P_{n-j}e_{k_j}$ for $j = 1, \cdots, n-i-1$. Since each $k_j > i$,

$$\begin{aligned} L_kP_{i+1}P_{i+2}\cdots P_{n-1}e_k &= L_kP_{i+1}P_{i+2}\cdots P_{n-2}e_{k_2} \\ &= \cdots \\ &= L_kP_{i+1}e_{k_{n-i-1}} = e_{k_{n-i-1}}. \end{aligned}$$

Now $P_{n-j}e_{k_{j+1}} = e_{k_j}$ implies $\tilde{L}_i e_k = e_k$.
(continued)

Proof

Since P_j permute $e_j \leftrightarrow e_k$ for some $k > j$, $P_j e_i = e_i$ for all $j > i$. Thus

$$\begin{aligned}\tilde{L}_i e_i &= P_{n-1}P_{n-2}\cdots P_{i+1}L_i e_i \\ &= P_{n-1}P_{n-2}\cdots P_{i+1}\left(\sum_{k=i}^{n} l_{ik}e_k\right) \\ &= e_i + \sum_{k=i+1}^{n} l_{ik}P_{n-1}P_{n-2}\cdots P_{i+1}e_k\end{aligned}$$

Since $P_{n-1}P_{n-2}\cdots P_{i+1}e_k = e_j$ for some $j > i$, $\tilde{L}_i e_i = (0, \cdots, 1, \cdots)$. Hence $\tilde{L}_i$ is a unit lower triangular matrix.

Matrix norm

A matrix norm $\|\cdot\|$ satisfies

- $\|\cdot\|$ is a norm on $\mathbb{R}^{m\times n}$
- $\|AB\| \leq \|A\|\|B\|$ for all $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times p}$.

Induced norm

Let $\|\cdot\|_n$ and $\|\cdot\|_m$ be norms on $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Then the induced norm $\|\cdot\|$ on $\mathbb{R}^{m\times n}$ given by

$$\|A\| = \{\|Ax\|_m : x \in \mathbb{R}^n, \|x\|_n \leq 1\}$$

is a norm. If we define $\|\cdot\|$ for all m, n, $\|\cdot\|$ is a matrix norm.

Euclidean norm

Suppose we give the Euclidean norm on $\mathbb{R}^n$ for all n. Let $\|\cdot\|_2$ the induced matrix norm. For $A \in \mathbb{R}^{m \times n}$, let σ_1 be the largest singular value of A. Then

$$\|A\|_2 = \sigma_1.$$

The QR factorization

Suppose $A \in \mathbb{R}^{m\times n}$ has full rank. Then A has the QR factorization where $Q \in \mathbb{R}^{m\times m}$ is orthogonal and $R \in \mathbb{R}^{m\times n}$ is upper triangular.

Householder Transformation

Let $x, y \in \mathbb{R}^n$ be such that $x \neq y$ and $\|x\|_2 = \|y\|_2$. Then there is a orthogonal matrix $U \in \mathbb{R}^{n\times n}$ such that $Ux = y$. This U is given by $U = I - 2u \otimes u$ where $u = (x - y)/\|x - y\|_2$.

Computing QR

Using Householder Transformation, we can compute the QR factorization as follows: Let $v_1 = A_1$ and $\alpha_1 = -\operatorname{sgn}(v_1)\|v_1\|$. Define $x = \alpha_1 e_1 - v_1$ and $u_1 = x_1/\|x_1\|_2$. Compute $Q_1 = I_m - 2u_1 \otimes u_1$ and $A^{(2)} = Q_1 A$.
Apply this process for the lower right hand submatrix $B^{(2)}$ of $A^{(2)}$. Then we have $Q_1, \cdots, Q_n$ and $Q_n Q_{n-1} \cdots Q_1 A = R$.

Finite dimensional

- Every norm on $\mathbb{R}^n$ is equivalent, i.e. given two norms $\|\cdot\|, \|\cdot\|_*$, there are $c_1, c_2 > 0$ such that

$$c_1\|x\|_* \leq \|x\| \leq c_2\|x\|_*$$

- $\mathbb{R}^n$ is complete under any norm.

Infinite dimensional

- l^2 is an infinite inner product space over $\mathbb{R}$.
- In l^2, the Bolzano-Weierstrass theorem fails.
- $C[a, b]$ is complete under $L^\infty[a, b]$ norm but is not complete under $L^2(a, b)$ norm.
  - $\{x^k\}$ on $[0, 1]$ is a Cauchy sequence under $L^2(0, 1)$ but not under $L^\infty[0, 1]$.

Funtional analysis

- For any normed vector space V, V^* is complete under the induced norm.
- Let H be a Hilbert space. If S is a closed subspace of H,
  - the projection theorem holds.
  - $S^{\perp\perp} = S$.
- For any $f \in H^*$, there is a unique $u \in H$ such that

$$f(v) = \langle v, u \rangle_H \text{ for all } v \in H.$$

Weak convergence

Let H be a Hilbert space over $\mathbb{R}$ and let $\{x_k\}$ be a sequence in H.

- Then $x_k \to x$ weakly iff $\langle x_k, u \rangle_H \to \langle x, u \rangle_H$ for all $u \in H$.
- If $x_k \to x$ weakly and $\|x_k\| \to \|x\|$, then $x_k \to x$ strongly.

The End