Class #3

Tuesday, January 28, 2025 10:20 AM

Variants of LS

Weighted LS

$$f(x) = \sum_{i=1}^{k} w_i (a_i^T x - b_i)^2 \qquad w_i \geq 0$$

$$= \| W(Ax - b) \|_2^2 \text{ , where}$$

$$W = \text{diag}(w_1, w_2, --, w_k)$$

$$= \| \underbrace{WAx}_{\tilde{A}} - \underbrace{Wb}_{\tilde{b}} \|_2^2$$

Regularized LS

$$f(x) = \|Ax - b\|_2^2 + \rho \underbrace{\|x\|_2^2}_{\text{regularizer}}$$

ρ → regularization parameter $\rho \geq 0$

$f(x)$ → Ridge regression

Linear Programming

$$\min_x \underbrace{c^T x} \quad f_0(x) \text{ is linear}$$

$$\text{subject to} \quad \underbrace{a_i^T x} \leq b_i \text{ , } i = 1, --, m$$

$f_1(x) --, f_m(x)$ linear

Typical

$f_1(x), \ldots, f_m(x)$ linear

Typical

$O(n^2 m)$ complexity if $m \geq n$

$$\underbrace{\min_{x} \max_{i=1,\ldots,k} |a_i^T x - b_i|}_{\text{chebyshev approximation}} = \min_{x} \|Ax - b\|_\infty$$

Compare LS

$$\min_{x} \|Ax - b\|_2^2$$

Review of key linear algebra concepts

Inner product

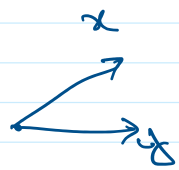
Given $x, y \in \mathbb{R}^n$, the standard inner product is

$$\langle x, y \rangle = x^T y = \sum_{i=1}^{n} x_i y_i$$

and $x^T x = \|x\|_2^2$, where $\|\cdot\|_2$ is Euclidean norm

Euclidean norm

Angle between two vectors: θ



$$x^T y = \|x\|_2 \|y\|_2 \cos(\theta)$$

If $x^T y = 0 \iff x \perp y$

If $x^T y > 0 \iff x$ and y make an acute angle

Euclidean norm: $\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$

Cauchy-Schwarz Inequality

$$|x^T y| \leq \|x\|_2 \|y\|_2$$

Inner Product between matrices

Given $X, Y \in \mathbb{R}^{m \times n}$, inner product on matrices

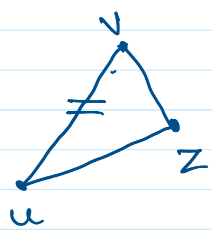
$$\langle X, Y \rangle = \operatorname{trace}(X^T Y) = \sum_{i=1}^{m} \sum_{j=1}^{n} X_{ij} Y_{ij}$$

$$\langle X, X \rangle = \|X\|_F^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} X_{ij}^2$$

What is a norm? Focus on $\mathbb{R}^n$

A function $f : \mathbb{R}^n \to \mathbb{R}$ with $\operatorname{dom} f = \mathbb{R}^n$ is called a norm if:

- $f(x) \geq 0 \quad \forall x \in \mathbb{R}^n$ — non-negativity
- $f(x) = 0$ only if $x = 0$ — definiteness
- $f(tx) = |t| f(x) \quad \forall t \in \mathbb{R}$, — homogeneity
- $f(x+y) \leq f(x) + f(y)$, — Triangle inequality



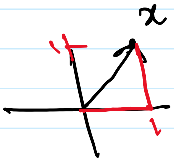
$\|x\|$

$\|x\|_2 \Rightarrow$ Euclidean norm $\Rightarrow$ l_2-norm

l_p-norm, $p \geq 1$

$$\|x\|_p = \left(|x_1|^p + |x_2|^p + \cdots + |x_n|^p\right)^{1/p}$$

l_1-norm $\Rightarrow \|x\|_1 = \sum_{i=1}^{n} |x_i|$



$\|x\|_1 = 2$

$\|x\|_2 = \sqrt{2}$

$\|x\|_\infty = 1$

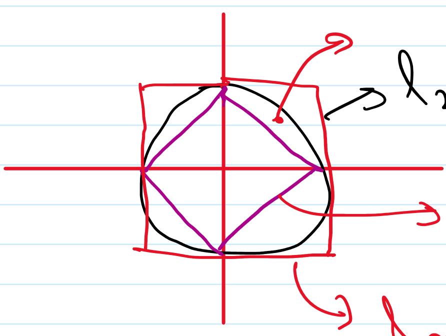
$p \to \infty$

l_∞-norm $\Rightarrow \|x\|_\infty = \max_{i=1,\dots,n} |x_i|$

↳ Chebyshev norm

<u>Unit ball in $\mathbb{R}^n$</u>

$$B = \{x \in \mathbb{R}^n,\ \|x\|_p \leq 1\}$$



All norms induce distances (metrics) on the vector space.

$$\text{dist}(x,y) = \|x-y\|_p$$

- $\text{dist}(x,y) = \text{dist}(y,x)$
- $\text{dist}(x,y) \geq 0 \quad \forall\ x,y$
- $\text{dist}(x,y) = 0$ only if $x=y$
- $\text{dist}(x,y) \leq \text{dist}(x,z) + \text{dist}(z,y)$

Space of Symmetric matrices : $S^n \Rightarrow n\times n$ Symmetric matrices

Eigenvalue Decomposition

A matrix $M \in \mathbb{R}^{n\times n}$ Said to have an eigenvalue decomposition if

$$M = \underline{U \Lambda U^{-1}} \Rightarrow M^T = U^{-T} \underline{\Lambda^T} U^T$$

$$= U^{-T} \Lambda U^T$$

wher $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$

$\lambda_1, \ldots, \lambda_n \Rightarrow$ eigenvalues of M

Columns of $U \Rightarrow$ eigenvectors of $M \Rightarrow u_1, \ldots, u_n$

If u_i is an eigenvector $\Leftrightarrow$

$$Mu_i = \lambda_i u_i$$

Let $M \in S^n \Rightarrow$ Then M always has an EVD and the eigenvalues are always real.

$$M^T = M \Leftrightarrow U \Lambda U^T = M$$

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

The eigenvectors are orthogonal to each other.
They are basis of $\mathbb{R}^n$.

$S_+^n \subset S^n$: S_+^n : Positive Semidefinite symmetric matrices $\hookrightarrow$ PSD

$\hookrightarrow \lambda_1, \ldots, \lambda_n \geq 0$

$S_{++}^n \subset S^n$: S_{++}^n : Positive definite symmetric matrices

$\lambda_1, \lambda_2, \ldots, \lambda_n > 0$ $\hookrightarrow$ PD

If $P \in S_+^n \iff P \succeq 0$

$P \in S_{++}^n \iff P \succ 0$

If $P \in S_+^n \iff x^T P x \geq 0 \quad \forall\, x \in \mathbb{R}^n$

If $P \in S_{++}^n \iff x^T P x > 0 \quad \forall\, x \in \mathbb{R}^n$

If $P \in S_+^n \Rightarrow$ Define $P^{1/2}$

$$P^{1/2} = U \Lambda^{1/2} U^T$$

$\hookrightarrow \mathrm{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n})$

$$P^{1/2} P^{1/2} = P$$

$$(U \Lambda^{1/2} U^T)(U \Lambda^{1/2} U^T)$$

$$\underbrace{(U\Lambda^{1/2}U^T)(U\Lambda^{1/2}U^T)}_{=I}$$

$$U\Lambda^{1/2}\Lambda^{1/2}U^T = U\Lambda U^T = P$$

P-Quadratic norm

Let $P \in S^n_+$

$$\|x\|_P = \underbrace{(x^T P x)^{1/2}}_{(x^T x)^{1/2} = \|x\|_2} = \|P^{1/2}x\|_2$$

$P = I$

$$\Leftrightarrow \|x\|_2 = \|x\|_I$$

$P \succ Q \Leftrightarrow P - Q \succ 0 \Rightarrow$ Notation

Norms on matrices $\Longrightarrow X \in \mathbb{R}^{m \times n}$

$$\|X\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} X_{ij}^2}$$

$$\|X\|_{sav} = \sum_{i=1}^{m}\sum_{j=1}^{n} |X_{ij}|$$

$$\|X\|_{max} = \max\{|X_{ij}|,\ i=1,\dots,m;\ j=1,\dots,n\}$$

All norms in finite-dimensional spaces (in particular, $\mathbb{R}^n$)

All norms in finite-dimensional spaces (in particular, $\mathbb{R}^n$) are "equivalent".

There always exists scalars α and β (which may depend on 'n') for given norms $\|\cdot\|_a$ and $\|\cdot\|_b$ s.t.

$$\alpha \|x\|_a \le \|x\|_b \le \beta \|x\|_a$$

e.g., $\|x\|_1 \le n \|x\|_\infty$

$\hookrightarrow \beta = n$