Convex Optimization: Reading Notes 1

GKxx

May 6, 2022

1 Least-squares problems

A least-squares problems is an optimization problem of the form

$$\min \quad f(x) = \|Ax - b\|_2^2 = \sum_{i=1}^{k} (a_i^T x - b_i)^2,$$

where $A \in \mathbb{R}^{k\times n}$ with $k \geqslant n$, and a_i^T are the rows of A, and the vector $x \in \mathbb{R}^n$ is the optimization variable. The gradient of the objective function is

$$\nabla f(x) = \nabla \left((Ax - b)^T(Ax - b)\right) = 2A^TAx - 2A^Tb.$$

Setting this to zero yields $A^TAx = A^Tb$. So the analytical solution to the least-squares problem, if A^TA is invertible, is $x = (A^TA)^{-1}A^Tb$.

Definition 1.1 (Moore-Penrose generalized inverse)**.** *For $A \in \mathbb{R}^{k\times n}$, consider the singular value decomposition $A = U\Sigma V^T$, as well as the thin SVD $A = \bar{U}\bar{\Sigma}\bar{V}^T$. Define $A^\dagger = \bar{V}\bar{\Sigma}^{-1}\bar{U}^T$, or $A^\dagger = V\Sigma^\dagger U^T$, where $\Sigma^\dagger$ is obtained from Σ by first replacing each nonzero singular value with its inverse and then transposing. The matrix $A^\dagger$ is called the* Moore-Penrose generalized inverse*.*

Remark 1.2. *In the SVD of $A \in \mathbb{R}^{k\times n}$, $U = \begin{bmatrix}\bar{U} & \tilde{U}\end{bmatrix}$ and $V = \begin{bmatrix}\bar{V} & \tilde{V}\end{bmatrix}$, where $\bar{U}, \tilde{U}, \bar{V}$ and $\tilde{V}$ are the orthonormal bases of $\mathcal{R}(A), \mathcal{N}\left(A^T\right), \mathcal{R}\left(A^T\right)$ and $\mathcal{N}(A)$, respectively.*

Lemma 1.3. $\mathcal{R}\left(A^\dagger\right) = \mathcal{R}\left(A^T\right)$.

Proof. $A^\dagger = \bar{V}\bar{\Sigma}^{-1}\bar{U}^T \Rightarrow \mathcal{R}\left(A^\dagger\right) \subseteq \mathcal{R}\left(\bar{V}\right)$. Moreover, $\bar{V} = A^\dagger\bar{U}\bar{\Sigma} \Rightarrow \mathcal{R}\left(\bar{V}\right) \subseteq \bar{A}^\dagger$. Therefore $\mathcal{R}\left(A^\dagger\right) = \mathcal{R}\left(\bar{V}\right)$. Since $\bar{V}$ is the orthonormal basis of $\mathcal{R}\left(A^T\right)$, we obtain $\mathcal{R}\left(A^\dagger\right) = \mathcal{R}\left(A^T\right)$. $\square$

Lemma 1.4. *If the linear system of equations $Ax = b$ is consistent, then $A^\dagger b$ is the unique solution of minimal Euclidean norm.*

Proof. First check that $A\left(A^\dagger b\right) = b$. Since the system is consistent, $b \in \mathcal{R}(A)$ and we have that

$$A\left(A^\dagger b\right) = \bar{U}\bar{\Sigma}\bar{V}^T\bar{V}\bar{\Sigma}^{-1}\bar{U}^Tb = \bar{U}\bar{U}^Tb = \Pi_{\mathcal{R}(A),\mathcal{R}(A)^\perp}(b) = b,$$

where $\Pi_{S,T}(x)$ is the projection of x onto S along T. Now let $A^\dagger b + \xi$ be any other solution, where $\xi \in \mathcal{N}(A)$. We have that

$$\left|\left|A^\dagger b + \xi\right|\right|_2^2 = \left(A^\dagger b + \xi\right)^T \left(A^\dagger b + \xi\right) = \left|\left|A^\dagger b\right|\right|_2^2 + \|\xi\|_2^2 + 2\xi^T A^\dagger b.$$

Here $A^\dagger b \in \mathcal{R}\left(A^\dagger\right) = \mathcal{R}\left(A^T\right)$, while $\xi \in \mathcal{N}(A) = \mathcal{R}\left(A^T\right)^\perp$, so $\xi^T A^\dagger b = 0$. Therefore

$$\left|\left|A^\dagger b + \xi\right|\right|_2^2 = \left|\left|A^\dagger b\right|\right|_2^2 + \|\xi\|_2^2 > \left|\left|A^\dagger b\right|\right|_2^2.$$

$\square$

Proposition 1.5. *The unique solution of minimal Euclidean norm to the least-squares problem is* $A^\dagger b$.

Proof. We have known that the solution to the least-squares problem must satisfy $A^T A x = A^T b$. Note that $A^T A = \bar{V}\bar{\Sigma}\bar{U}^T\bar{U}\bar{\Sigma}\bar{V}^T = \bar{V}\bar{\Sigma}^2\bar{V}^T$, so the Moore-Penrose generalized inverse of $A^T A$ is

$$\left(A^T A\right)^\dagger = \bar{V}\bar{\Sigma}^{-2}\bar{V}^T.$$

From Lemma 1.4 we know that the unique solution of minimal Euclidean norm is

$$\left(A^T A\right)^\dagger A^T b = \bar{V}\bar{\Sigma}^{-2}\bar{V}^T\bar{V}\bar{\Sigma}\bar{U}^T b = \bar{V}\bar{\Sigma}^{-1}\bar{U}^T b = A^\dagger b.$$

$\square$

2 Affine sets, convex sets and cones

Definition 2.1 (Affine set)**.** *A set* $C \subseteq \mathbb{R}^n$ *is* affine *if for any* $x_1, x_2 \in C$ *and* $\theta \in \mathbb{R}$, *we have* $\theta x_1 + (1-\theta)x_2 \in C$.

Definition 2.2 (Affine combination)**.** *The* affine combination *of* k *points* $x_1, \cdots, x_k$ *is the linear combination*

$$\theta_1 x_1 + \cdots + \theta_k x_k,$$

where $\theta_1 + \cdots + \theta_k = 1$.

Proposition 2.3. *The solution set of a system of linear equations is an affine set. Conversely, every affine set can be expressed as the solution set of a system of linear equations.*

Proof. We only show the converse. Suppose C is an affine set. If $C = \varnothing$ this is trivial. When $C \neq \varnothing$, take any $x_0 \in C$ and the set $V = C - x_0 = \{x - x_0 \mid x \in C\}$ is a linear subspace. Suppose $\dim V^\perp = d$ and $B = \{b_1, \cdots, b_d\}$ is a basis for $V^\perp$. Then

$$x \in V \iff \forall \xi \in V^\perp, x \perp \xi \iff \forall i \in [d], x \perp b_i.$$

Now define the matrix $A = \begin{bmatrix} b_1 & \cdots & b_d \end{bmatrix}^T$, the i-th row of which is b_i^T, and then $x \in V \iff Ax = 0$. Therefore $x \in C \iff Ax = Ax_0$, which shows that C is the solution set of $Ax = Ax_0$. $\square$

Definition 2.4 (Affine hull)**.** *For a set $C \subseteq \mathbb{R}^n$, the set of all affine combinations of points in C is called the* affine hull *of C, denoted $\mathbf{aff}\, C$:*

$$\mathbf{aff}\, C = \{\theta_1 x_1 + \cdots + \theta_k x_k \mid x_1, \cdots, x_k \in C, \theta_1 + \cdots + \theta_k = 1\}.$$

Remark 2.5. *$\mathbf{aff}\, C$ is the smallest affine set that contains C. For any affine set $S \supseteq C$, $\mathbf{aff}\, C \subseteq S$.*

Definition 2.6 (Relative interior)**.** *The* relative interior *of a set C, denoted $\mathbf{relint}\, C$, is defined to be its interior relative to $\mathbf{aff}\, C$, i.e.*

$$\mathbf{relint}\, C = \{x \in C \mid \exists r > 0 \text{ s.t. } B(x,r) \cap \mathbf{aff}\, C \subseteq C\}.$$

Here $B(x,r)$ is the ball of radius r and center x defined by any norm.

Remark 2.7 (Interior)**.** *The* interior *of a set C*

$$\{x \in C \mid \exists r > 0 \text{ s.t. } B(x,r) \subseteq C\}$$

is contained in the set $\mathbf{relint}\, C$.

Proposition 2.8. *All norms define the same relative interior.*

Proof Sketch. We can show that there exists nonnegative constants c and d such that

$$c\,\|x\| \leqslant \|x\|_1 \leqslant d\,\|x\|$$

holds for every $x \in \mathbb{R}^n$, no matter what norm $\|\cdot\|$ is chosen. Here $\|\cdot\|_1$ is the ℓ_1-norm. In that sense, all norms are equivalent in a vector space. □

Definition 2.9 (Cone)**.** *A set $C \subseteq \mathbb{R}^n$ is a* cone *if $\theta x \in C$ for every $x \in C$ and $\theta \geqslant 0$.*

Definition 2.10 (Conic combination)**.** *The* conic combination *of k points $x_1, \cdots, x_k \in \mathbb{R}^n$ is of the form*

$$\theta_1 x_1 + \cdots + \theta_k x_k,$$

where $\theta_1, \cdots, \theta_k \geqslant 0$.

Remark 2.11. *The* convex combination *is both a conic combination and an affine combination.*

Definition 2.12 (Convex cone)**.** *The set of* convex cones *is the intersection of the set of convex sets and that of cones.*

Proposition 2.13. *C is a convex cone if and only if C contains all conic combinations of points in itself.*

Proof. $\Leftarrow$: Suppose C contains all conic combinations of points in itself, i.e. for every $x_1, x_2 \in C$ and $\theta_1, \theta_2 \geqslant 0$ we have $\theta_1 x_1 + \theta_2 x_2 \in C$. Set $\theta_2 = 0$ and we obtain that $\forall \theta_1 \geqslant 0, \theta_1 x_1 \in C$, so C is a cone. Setting $\theta_1 + \theta > 0$ yields

$$\frac{\theta_1}{\theta_1 + \theta_2} x_1 + \frac{\theta_2}{\theta_1 + \theta_2} x_2 = \frac{1}{\theta_1 + \theta_2} (\theta_1 x_1 + \theta_2 x_2) \in C,$$

which means that C is a convex set. Hence C is a convex cone.

$\Rightarrow$: For a convex cone C, suppose $x_0 = \theta_1 x_1 + \cdots + \theta_k x_k$ is a conic combination of k points $x_1, \cdots, x_k \in C$, where $\theta_1, \cdots, \theta_k \geqslant 0$. If $\Theta = \theta_1 + \cdots + \theta_k = 0$, then all the θ_i's are zero and $x_0 = 0 \in C$ is obvious. When $\Theta > 0$, we know that

$$\frac{1}{\Theta} x_0 = \frac{\theta_1}{\Theta} x_1 + \cdots + \frac{\theta_k}{\Theta} x_k \in C$$

since C is convex, noting that

$$\frac{\theta_1}{\Theta} + \cdots + \frac{\theta_k}{\Theta} = 1.$$

Since C is a cone, x_0 is also contained in C. □

3 Balls, ellipsoids and norm cones

Definition 3.1 (Norm ball). *A* norm ball *in* $\mathbb{R}^n$ *has the form*

$$B(x_c, r) = \{x \mid \|x - x_c\| \leqslant r\},$$

where $r > 0$. *Here* $\|\cdot\|$ *is any norm.*

Definition 3.2 (Euclidean ball). *A* Euclidean ball *is a norm ball defined by the Euclidean norm* $\|\cdot\|_2$.

Remark 3.3. $B(x_c, r) = \{x_c + ru \mid \|u\|_2 \leqslant 1\}$.

Proposition 3.4. *A norm ball is convex.*

Proof. Suppose $x_1, x_2 \in B(x_c, r)$. Then for any $0 \leqslant \theta \leqslant 1$,

$$\begin{aligned} \|\theta x_1 + (1-\theta) x_2 - x_c\| &= \|\theta(x_1 - x_c) + (1-\theta)(x_2 - x_c)\| \\ &\leqslant \theta \|x_1 - x_c\| + (1-\theta) \|x_2 - x_c\| \\ &\leqslant \theta r + (1-\theta) r = r. \end{aligned}$$

Therefore the convex combination $\theta x_1 + (1-\theta) x_2$ is also contained in $B(x_c, r)$. □

Definition 3.5 (Euclidean ellipsoid). *A* Euclidean ellipsoid *in* $\mathbb{R}^n$ *has the form*

$$\mathcal{E}(x_c, P) = \left\{x \mid (x - x_c)^T P^{-1} (x - x_c) \leqslant 1\right\},$$

where $x_c \in \mathbb{R}^n$ *is its* center *and* $P \in \mathbb{S}^n_{++}$ *determines how far the ellipsoid extends in every direction from* x_c. *The lengths of the semi-axes of* $\mathcal{E}(x_c, P)$ *are given by* $\sqrt{\lambda_i}$, *where* λ_i*'s are the eigenvalues of* P.

Remark 3.6. *For* $r > 0$, $\mathcal{E}(x_c, r^2 I)$ *is the Euclidean ball* $B(x_c, r)$.

Remark 3.7. $\mathcal{E}(x_c, P) = \{x_c + Au \mid \|u\|_2 \leqslant 1\}$, *where* $A = P^{1/2}$ *is defined by first diagonalizing* $P = Q\Lambda Q^T$ *and then taking* $A = Q\Lambda^{1/2}Q^T$. A *is also symmetric and positive definite.*

Proposition 3.8. *A Euclidean ellipsoid is convex.*

Proof. For a Euclidean ellipsoid $\mathcal{E} = \{x_c + Au \mid \|u\|_2 \leqslant 1\}$, consider any $x_c + Au_1, x_c + Au_2 \in \mathcal{E}$ and $\theta \in [0, 1]$, the convex combination induced by which satisfies

$$\theta(x_c + Au_1) + (1 - \theta)(x_c + Au_2) = x_c + A(\theta u_1 + (1 - \theta)u_2),$$

and by triangle inequality we have

$$\|\theta u_1 + (1 - \theta)u_2\|_2 \leqslant \theta \|u_1\|_2 + (1 - \theta) \|u_2\|_2 \leqslant 1.$$

Therefore the convex combination is contained in $\mathcal{E}$. □

Proof by convex function. For a Euclidean ellipsoid $\mathcal{E}(x_c, P)$, consider the function $f(x) = (x - x_c)^T P^{-1}(x - x_c)$. f is convex because $\nabla^2 f(x) = 2P^{-1} \succ 0$. Then for every $x_1, x_2 \in \mathcal{E}(x_c, P)$ and every $\theta \in [0, 1]$, we have

$$\begin{aligned} f(\theta x_1 + (1 - \theta)x_2) &\leqslant \theta f(x_1) + (1 - \theta)f(x_2) \\ &\leqslant \theta + (1 - \theta) = 1. \end{aligned}$$

Therefore the convex combination $\theta x_1 + (1 - \theta)x_2$ is contained in $\mathcal{E}(x_c, P)$, so $\mathcal{E}(x_c, P)$ is convex. □

Question 3.9. *The inequality*

$$(x - x_c)^T P^{-1}(x - x_c) \leqslant 1$$

can be rewritten as

$$1 - (x - x_c)^T P^{-1}(x - x_c) \geqslant 0,$$

the left-hand side of which is the Schur complement of *the matrix*

$$S = \begin{bmatrix} P & x - x_c \\ (x - x_c)^T & 1 \end{bmatrix},$$

so the inequality holds if and only if $S \in \mathbb{S}^n_+$. *What can we obtain with this?*

Definition 3.10 (Norm cone)**.** *The* norm cone *associated with the norm* $\|\cdot\|$ *is the set*

$$C = \{(x, t) \mid \|x\| \leqslant t\} \subseteq \mathbb{R}^{n+1}.$$

Proposition 3.11. *A norm cone is a convex cone.*

Proof. We will show the convexity of a norm cone $C = \{(x, t) \mid \|x\| \leqslant t\}$. For every (x_1, t_1), $(x_2, t_2) \in C$ and $\theta \in [0, 1]$,

$$\|\theta x_1 + (1 - \theta)x_2\| \leqslant \theta \|x_1\| + (1 - \theta) \|x_2\| \leqslant \theta t_1 + (1 - \theta)t_2.$$

Therefore the convex combination $\theta(x_1, t_1) + (1 - \theta)(x_2, t_2) \in C$. □

Definition 3.12 (Second-order cone)**.** *A* second-order cone *is the norm cone defined by the Euclidean norm. It is also called the* quadratic cone*, the* Lorentz cone *or the* ice-cream cone*.*

Definition 3.13 (Positive semidefinite cone)**.** *The set $\mathbb{S}^n_+$, which denotes the set of all positive semidefinite matrices, induces a* positive semidefinite cone*.*

Proposition 3.14. *$\mathbb{S}^n_+$ is a convex cone.*

Proof. Since for any $A \in \mathbb{S}^n_+$ and any $\theta \geqslant 0$, θA is also positive semidefinite, $\mathbb{S}^n_+$ is a cone. Moreover, for any $\theta \in [0, 1]$ and every $A, B \in \mathbb{S}^b_+$,

$$x^T (\theta A + (1 - \theta)B)\, x = \theta x^T A x + (1 - \theta) x^T B x \geqslant 0$$

holds for every vector x, so $\theta A + (1 - \theta)B \in \mathbb{S}^n_+$. This shows the convexity. □