Lecture #16

Geometric Problems

Projection on a set:

Suppose $\underbrace{\mathcal{C}}_{\text{closed set}} \subseteq \mathbb{R}^n$.

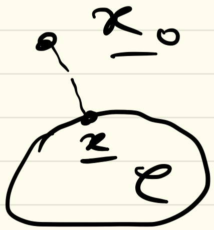
$$\underline{x}^{opt} = \underset{\underline{x} \in \mathcal{C}}{\operatorname{argmin}} \|\underline{x} - \underline{x}_0\| = \operatorname{proj}^{\|\cdot\|}_{\mathcal{C}}(\underline{x}_0)$$

Given a point $\underline{x}_0 \in \mathbb{R}^n$, project it on $\mathcal{C}$.

$\Leftrightarrow$ Find the point in $\mathcal{C}$ that is nearest to $\underline{x}_0$

$$\underset{\underline{x} \in \mathbb{R}^n}{\text{minimize}} \; \|\underline{x} - \underline{x}_0\|$$

$$\text{s.t.} \quad \underline{x} \in \mathcal{C}$$



$$\underline{x}^{opt} = \text{proj}_{\mathcal{C}}^{\|\cdot\|}(\underline{x}_0)$$

closest point of $\underline{x}_0$ in $\mathcal{C}$.

- If $\mathcal{C}$ is closed & convex, & $\|\cdot\|$ is strictly convex, then $\underline{x}^{opt}$ is unique.

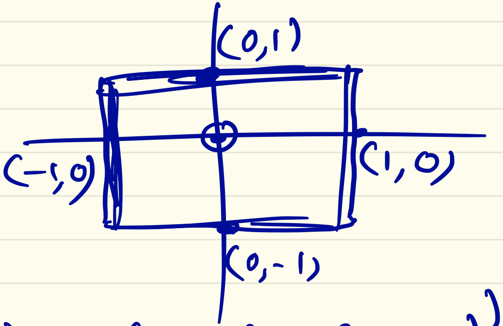
- Interesting converse:

If for every $\underline{x}_0$, $\exists$ unique Euclidean projection of $\underline{x}_0$ on $\mathcal{C}$, then $\mathcal{C}$ is closed & convex.

$$\mathrm{proj}_{\mathcal{C}}^{\|\cdot\|}(x_0) = \underset{\underline{x} \in \mathcal{C}}{\mathrm{argmin}} \ \|\underline{x} - \underline{x}_0\|$$

Examples :

① $\mathcal{C}$ = square in $\mathbb{R}^2$
bndy.



$\underline{x}_0 = \underline{0}$

$$\mathrm{proj}_{\mathcal{C}}^{\|\cdot\|_1}(\underline{0}) = \{(0,1), (0,-1), (1,0), (-1,0)\}$$

$$\mathrm{proj}_{\mathcal{C}}^{\|\cdot\|_\infty}(\underline{0}) = \{\text{all pt-s on the bndy}\}$$

Example 2

Let $\mathcal{C} = \{ X \in \mathbb{R}^{m \times n} \mid \text{rank}(X) \leq k \}$

where $k \leq \min\{m, n\}$

Let $X_0 \in \mathbb{R}^{m \times n}$

$$\text{proj}_{\mathcal{C}}^{\|\cdot\|_2}(X_0) = Y = \sum_{i=1}^{\min\{k, r\}} \sigma_i u_i v_i^T$$

$$\underbrace{X_0 = \sum_{i=1}^{r} \sigma_i u_i v_i^T}_{r = \text{rank}(X_0)}$$

SVD of X_0

(singular value decomposⁿ)

$\|A\|_2$ norm = spectral norm

$$= \sqrt{\lambda_{max}(A^T A)}$$

$$= \sigma_{max}(A)$$

Euclidean projection on $\mathcal{P}$ (polyhedron)

$$\left.\begin{aligned} \mathrm{proj}_{\mathcal{P}}^{\|\cdot\|_2}(\underline{x}_0) \quad \operatorname{argmin} \; & \|\underline{x} - \underline{x}_0\|_2^2 \\ \text{s.t.} \quad & A\underline{x} \preceq \underline{b} \end{aligned}\right\} \text{QP}$$

Special cases may have analytical sol^n

- $\underbrace{\mathcal{C} = \{\underline{x} \in \mathbb{R}^n \mid \underline{a}^T\underline{x} = b\}}_{\text{hyperplane}}$

$$\mathrm{proj}_{\mathcal{C}}^{\|\cdot\|_2}(\underline{x}_0) = \underline{x}_0 + \frac{(b - \underline{a}^T\underline{x}_0)\,\underline{a}}{\|\underline{a}\|_2^2}$$

- $\mathcal{C} = \{ \underline{x} \in \mathbb{R}^n \mid \underline{a}^T \underline{x} \leq b \}$

Halfspace

$$\mathrm{proj}_{\mathcal{C}}^{\|\cdot\|_2}(\underline{x}_0) = \begin{cases} \underline{x}_0 + \dfrac{(b - \underline{a}^T \underline{x}_0)\,\underline{a}}{\|\underline{a}\|_2^2} & \text{if } \underline{a}^T \underline{x}_0 > b \\ x_0 & \text{if } \underline{a}^T \underline{x}_0 \leq b \end{cases}$$

- $\mathcal{C} = \{ \underline{x} \mid \underline{l} \leq \underline{x} \leq \underline{u} \}$

rectangle

$$\left\{ \mathrm{proj}_{\mathcal{C}}^{\|\cdot\|_2}(\underline{x}_0) \right\}_k = \begin{cases} l_k & \text{if } x_{0k} \leq l_k \\ x_{0k} & \text{if } l_k \leq x_{0k} \leq u_k \\ u_k & \text{if } x_{0k} \geq u_k \end{cases}$$

- $\|\cdot\|_2 \leftarrow$ Euclidean distance

$\mathcal{C} \equiv K$ (proper convex cone)

$$\text{proj}_K^{\|\cdot\|_2}(\underline{x}_0) = \underset{\underline{x} \in K}{\text{argmin}} \|\underline{x} - \underline{x}_0\|_2^2$$

$$\Updownarrow$$

$$\underline{x} \succcurlyeq_K \underline{0}$$

e.g. $K \equiv \mathbb{R}^n_{\geqslant 0}$ (non-negative orthant of $\mathbb{R}^n$)

$$\left\{\text{proj}_{K \equiv \mathbb{R}^n_{\geqslant 0}}^{\|\cdot\|_2}(\underline{x}_0)\right\}_k = \max\{x_{0k}, 0\}$$

zero-out any negative component.

- $K \equiv S^n_+$, $\|\cdot\|_F$

Let $X_0 \in S^n$

Both $\|\cdot\|_F$ and $\|\cdot\|_2$ give same answer

$$\text{proj}^{\|\cdot\|_F}_{S^n_+}(X_0) = \text{proj}^{\|\cdot\|_2}_{S^n_+}(X_0)$$

$$= \sum_{i=1}^{n} \max\{0, \lambda_i\} v_i v_i^T$$

= spectral decompos^n of X_0 & drop the negative eig. value terms

$\|A\|_F$ ← Frobenius norm

$$= \sqrt{\text{tr}(A^T A)}$$

$$\langle A, B\rangle_F = \text{tr}(A^T B)$$

$$\|A\|_F = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^2}$$

eig. value, eig. vector

$$X_0 = \underbrace{\sum_{i=1}^{n} \lambda_i v_i v_i^T}_{\text{spectral decompos}^n \text{ of } X_0}$$

- **Distance between 2 convex sets**

$$\mathcal{C} = \{ \underline{x} \in \mathbb{R}^n \mid f_i(\underline{x}) \le 0, \; i = 1, \dots, m \}$$

$$\mathcal{D} = \{ \underline{x} \in \mathbb{R}^n \mid g_i(\underline{x}) \le 0, \; i = 1, \dots, p \}$$

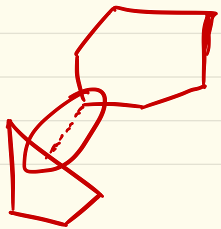
(*)
$$\mathrm{dist}(\mathcal{C}, \mathcal{D}) = \min_{\underline{x}, \underline{y} \in \mathbb{R}^n} \| \underline{x} - \underline{y} \|$$

$$\text{s.t.} \quad f_i(\underline{x}) \le 0, \; i = 1, \dots, m$$

$$g_i(\underline{y}) \le 0, \; i = 1, \dots, p$$

e.g. Distance bet^n 2 convex polyhedron

$$\left.\begin{aligned} &\min \; \| \underline{x} - \underline{y} \|_2^2 \\ &\text{s.t.} \; A_1 \underline{x} \le b_1 \\ &\qquad A_2 \cancel{\underline{x}}\, \underline{y} \le b_2 \end{aligned}\right\} \quad \underline{\text{QP}}$$



Dual of (*)

$$g(\underline{\lambda}, \underline{\mu}, \underline{\nu})$$

Lagrange multipliers for inequality constraints

$$= \inf_{\underline{x}, \underline{y}, \underline{w}} \left\{ \|\underline{w}\| + \sum_{i=1}^{m} \lambda_i f_i(\underline{x}) + \sum_{i=1}^{p} \mu_i g_i(\underline{y}) + \underline{\nu}^T(\underline{x} - \underline{y} - \underline{w}) \right\}$$

Rewrite (*):

$$\min_{\underline{x} \in \mathbb{R}^n,\ \underline{y} \in \mathbb{R}^n,\ \underline{w} \in \mathbb{R}^n} \|\underline{w}\|$$

s.t. $f_i(\underline{x}) \leq 0, \quad i = 1, \ldots, m$

$g_i(\underline{y}) \leq 0, \quad i = 1, \ldots, p$

$\underline{x} - \underline{y} = \underline{w}$

$$= \begin{cases} \boxed{\inf_{\underline{x}} \left\{ \sum_{i=1}^{m} \lambda_i f_i(\underline{x}) + \underline{\nu}^T \underline{x} \right\} + \inf_{\underline{y}} \left\{ \sum_{i=1}^{p} \mu_i g_i(\underline{y}) - \underline{\nu}^T \underline{y} \right\}} & \text{if } \|\underline{\nu}\|_* \leq 1 \\ = -\infty & \text{otherwise} \end{cases}$$

$$\Downarrow$$

Dual problem

$$\underset{\underline{\lambda}, \underline{\mu}, \underline{\nu}}{\text{maximize}} \quad \boxed{}$$

$$\text{s.t.} \quad \|\underline{\nu}\|_* \leq 1$$

$$\underline{\lambda} \succcurlyeq 0, \quad \underline{\mu} \succcurlyeq 0.$$

If $\underline{\lambda}, \underline{\mu}$ are dual feasible $(\underline{\lambda} \geq 0, \underline{\mu} \geq 0)$, then {inside the inf} $> 0 \ \forall \ \underline{x}, \underline{y}$

$\therefore$ For $\underline{x} \in \mathcal{C}$, $\underline{y} \in \mathcal{D}$

$$\underline{v}^T \underline{x} - \underline{v}^T \underline{y} > 0$$

(i.e.) $\underline{v}^T(\underline{x} - \underline{y}) > 0$

(i.e.) $\underline{v}$ defines a separating hyperplane that strictly separates $\mathcal{C}$ & $\mathcal{D}$

If strong duality holds, then following interpretation:

If dist. betn 2 convex sets > 0, then they are strictly separable by a hyp. plane.

e.g. suppose, $\mathcal{C} \equiv A_1 \underline{x} \preceq \underline{b}_1$

$\mathcal{D} \equiv A_2 \underline{x} \preceq \underline{b}_2$

Dual problem:

$$\max \left(-b_1^T \underline{\lambda} - b_2^T \underline{\mu} \right)$$

$$\text{s.t.} \quad A_1^T \underline{\lambda} + \underline{\nu} = 0$$

$$A_2^T \underline{\mu} - \underline{\nu} = 0$$

$$\| \underline{\nu} \|_* \leq 1$$

$$\underline{\lambda} \succeq 0, \quad \underline{\mu} \succeq 0$$

sep. hyp. plane:

$$\underline{\nu}^T (\underline{x} - \underline{y}) \geq -\underline{\lambda}^T b_1 - \underline{\mu}^T b_2 > 0$$

$\underline{\nu}$ defines separating hyp. plane.

Distance & Angle Problems:

Suppose $\underline{a}_1, \ldots, \underline{a}_n \in \mathbb{R}^n$ } ; n points in $\mathbb{R}^n$

$$\ell_1 := \|\underline{a}_1\|_2, \ldots, \ell_n := \|\underline{a}_n\|_2.$$

If $\underline{a}_1, \ldots, \underline{a}_n$ are linearly indep., then we call these pt.s a basis for $\mathbb{R}^n$

Length & angles can be expressed via Gram matrix:

$$\underset{n \times n}{A} = [\underline{a}_1 | \underline{a}_2 | \ldots | \underline{a}_n]$$

$$G := A^T A \succcurlyeq 0$$

(i.e.) $G(i,j) = \langle \underline{a}_i, \underline{a}_j \rangle$

$$G(i,i) = \ell_i^2$$

distance between pt.s $\underline{a}_i$ & $\underline{a}_j$:

$$d_{ij} = d(\underline{a}_i, \underline{a}_j) := \| \underline{a}_i - \underline{a}_j \|_2$$

$$= \left(l_i^2 + l_j^2 - 2 \langle \underline{a}_i, \underline{a}_j \rangle \right)^{1/2}$$

$$= \left\{ G(i,i) + G(j,j) - 2G(i,j) \right\}^{1/2}$$

Conversely,

$$G(i,j) = \frac{l_i^2 + l_j^2 - d_{ij}^2}{2}$$ affine in d_{ij}^2

Angle between $\underline{a}_i$ & $\underline{a}_j$, say θ_{ij} :

$$\theta_{ij} = \cos^{-1}\left(\frac{\langle \underline{a}_i, \underline{a}_j \rangle}{\| \underline{a}_i \|_2 \, \| \underline{a}_j \|_2} \right) = \cos^{-1}\left(\frac{G(i,j)}{l_i l_j} \right)$$

If $Q \in O(n)$ then

set of all $n \times n$ orthogonal matrices.

$Q\underline{a}_1, \ldots, Q\underline{a}_n$

has same Gram matrix G

$\Rightarrow$ same lengths & same angles

Relative distances & angles are invariant under pure rotation

$Q \in O(n)$

$\Updownarrow$

$QQ^T = Q^TQ = I_{n \times n}$

$\Updownarrow$

$Q^{-1} = Q^T$

$\Updownarrow$

$\det(Q) = \pm 1$

Notice $G := A^TA \in S_+^n$

$G \in S^n$ is Gram matrix of some pt.s $\underline{a}_1, \ldots, \underline{a}_n$

$\Updownarrow$

$G \succeq 0$, & has diag. $G(i,i) = l_i^2$

This implies that

a set of angle & distances are <u>realizable</u>

<u>if and only if</u> the associated Gram

matrix $G \succcurlyeq 0$ and $G(i,i) = l_i^2$