Projections onto different sets

6:56 PM

*Projections onto subspaces:

$\Pi_S(x) = \underset{y\in S}{\operatorname{argmin}} \underbrace{\|y-x\|}_{(\langle y-x,\, y-x\rangle)^{1/2}}$

closed, $\in$ inner product space

*One dimensional subspace:

$S_v = \{\lambda v : \lambda \in \mathbb{R}\}$ {one-dimensional sub-space generated by v}, $v \neq 0$

$\Pi_{S_v}(x) = \underset{y\in S_v}{\operatorname{argmin}} \|y-x\|$

{property of the projection}

- $(x - \Pi_{S_v}(x)) \perp v$

Proof: $\Pi_{S_v}(x) = \underset{y\in S_v}{\operatorname{argmin}} \|y-x\|^2$

now $\|y-x\|^2 = \|(y-x_v)-(x-x_v)\|^2$ // where x_v is choosen

$= \|y-x_v\|^2 + \|x-x_v\|^2 - 2\langle y-x_v, x-x_v\rangle$ // such than $\langle x-x_v, v\rangle = 0 \Leftrightarrow (x-x_v)\perp v$

// and $x_v \in S_v$, note that such a point will always exist as $x_v \in S_v \Rightarrow \exists\, x_v = \lambda v$

// $\langle x-x_v, v\rangle = \langle x, v\rangle - \langle x_v, v\rangle = \langle x,v\rangle - \lambda\langle v,v\rangle$, so by picking $\lambda = \frac{\langle x,v\rangle}{\|v\|^2}$ we can construct ($\langle v,v\rangle$: norm nonzero by definition $= \|v\|^2$)

$\langle y, x-x_v\rangle - \langle x_v, x-x_v\rangle$

$\lambda_1 v [\because y\in S_v]$, $\lambda v [\because x_v \in S_v]$

$= \lambda_1\langle v, x-x_v\rangle - \lambda\langle v, x-x_v\rangle = 0$

{$\langle x,y\rangle = \langle x,y\rangle$ by defn} // an $x_v = \lambda v$ such that $(x-x_v)\perp v$

$= \|y-x_v\|^2 + \|x-x_v\|^2$

Clearly this term will be minimized when $y = x_v$ // y is our decision variable

$= \frac{\langle x,v\rangle}{\|v\|^2} v$

$\therefore \boxed{\Pi_{S_v}(x) = \frac{\langle x,v\rangle}{\|v\|^2} v}$ {component of x along the direction v}

- Projection onto an arbitrary subspace: 'S'

* Theorem 2.1: $\{\mathcal{X}$ {inner product space}, $x\in\mathcal{X}$, S {subspace($\mathcal{X}$)}$\} \Rightarrow \exists!_{x^*\in S}\left(x^* = \underset{y\in S}{\operatorname{argmin}}\|y-x\| \Leftrightarrow (x-x^*)\perp S\right)$

* Proof: $\forall_{x\in\mathcal{X}}\ \exists_{u\in S}\ \exists_{z\in S^\perp}\ x = u+z$

// $\because$ By Orthogonal decomposition:

// $\{S$ {subspace($\mathcal{X}$)}$\}$ $\left(\forall_{x\in\mathcal{X}}\ \exists!_{(u,v)\in S\times S^\perp}\ x = u+v\right) \overset{\text{le}}{=} \mathcal{X} = S\oplus S^\perp$

{inner product space}

Boyd's version orthogonal decomposition law

$\left\{ \begin{array}{l} L: \text{subspace}(\mathcal{X}: \text{inner product space}) \\ L^\perp \text{ \{orthogonal complement of } L\} = \{y \mid \forall_{x\in L}\ y^T x\} \end{array}\right\}$

$\forall_{v\in\mathcal{X}}\ v = \Pi_L(v) + \Pi_{L^\perp}(v)$

$\forall_{x\in\mathcal{X}}\ \forall_{y\in S}\ \|y-x\|^2 = \|y-u-z\|^2 = \|y-u\|^2 + \|z\|^2 - 2\underbrace{\langle y-u, z\rangle}_{\shortparallel}$

$\langle y,z\rangle - \langle u,z\rangle = 0 - 0 = 0$

$\in S$, $\in S^\perp$, $\in S$, $\in S^\perp$

$= \|y-u\|^2 + \|z\|^2$

$\forall_{x\in\mathcal{X}}\ \forall_{y\in S}\ \|y-x\|^2 = \|y-u\|^2 + \|z\|^2$

$\forall_{x\in\mathcal{X}}\ \underset{y\in S}{\min}\|y-x\|^2 = \underset{y\in S}{\min}\left(\|y-u\|^2 + \|z\|^2\right) = \|z\|^2 + \underset{y\in S}{\min}\|y-u\|^2 = \|z\|^2$

0 for $y = u$

the pair (u,z) is uniquely determined by x, so they are constant w.r.t the optimization problem

$\therefore \Pi_S(x) = \underset{y\in S}{\operatorname{argmin}}\|y-x\| = u$

$\Leftrightarrow x - x^* = x - u \in S^\perp$ [$\because x = u + z$, $z \in S^\perp \Rightarrow z = x - u \in S^\perp$]

$\Leftrightarrow$ [By definition]

$(x - x^*) \perp S$ ∎

- Corollary 2.1: Projection on affine set

$\{\mathcal{X}$ {inner product space}, $x\in\mathcal{X}$, $A = x^{(0)} + S\} \Rightarrow \exists!_{x^*\in A}\left(x^* = \underset{y\in A}{\operatorname{argmin}}\|y-x\| \Leftrightarrow (x-x^*)\perp S\right)$

proof: $A = x^{(0)} + S$

$\Leftrightarrow \forall_y\ (y\in A \Leftrightarrow \exists_{z\in S}\ y = z + x^{(0)})$

$\therefore \underset{y\in A}{\min}\|y-x\|^2 = \underset{z\in S}{\min}\|z + x^{(0)} - x\|^2$

$= \underset{z\in S}{\min}\|z - (x - x^{(0)})\|^2 = \|s^\perp_{(x-x^{(0)})}\|^2$ where $\underbrace{x - x^{(0)}}_{\in\mathcal{X}} = \underbrace{s_{(x-x^{(0)})}}_{\in S} + \underbrace{s^\perp_{(x-x^{(0)})}}_{\in S^\perp}$ by Theorem 2.1

($\mathcal{X}$: inner product space; S: subspace of $\mathcal{X}$)

with $y = z^* = s_{(x-x^{(0)})}$

$\therefore$ in original variable, $y^* = z^* + x^{(0)} = s_{(x-x^{(0)})} + x^{(0)}$

so, $x^* = \underset{y\in A}{\operatorname{argmin}}\|y-x\| = s_{(x-x^{(0)})} + x^{(0)}$

$\rightarrow (x - x^*) = x - s_{(x-x^{(0)})} - x^{(0)} = s^\perp_{(x-x^{(0)})}$

$\Leftrightarrow (x - x^*) \perp S$ (proved)

*Euclidean Projection on a line:

$p\in\mathbb{R}^n$

$L = \{x_0 + \text{span}(u) : \|u\|_2 = 1\}$

by taking coefficient of u arbitrarily, we can construct line of any length.

$p^* = \underset{x\in L}{\operatorname{argmin}}\|x - p\|_2^2 = x_0 + \left(\underset{\lambda\in\mathbb{R}}{\operatorname{argmin}}\|x_0 + \lambda u - p\|_2^2\right) u = x_0 + u^T(p - x_0)u$

line origin; line spanner; the point whose projection we are taking by taking coefficient of u arbitrarily, we can construct line of any length.

line origin, line spanner, the point whose projection we are taking

$$p^* = \underset{x \in L}{\operatorname{argmin}} \|x - p\|_2^2 = x_0 + \left(\underset{\lambda \in \mathbb{R}}{\operatorname{argmin}} \|x_0 + \lambda u - p\|_2^2\right) u = x_0 + u^T(p - x_0) u$$

$x \in L \Leftrightarrow \exists_\lambda \; x = x_0 + \lambda u$ (given; given $\|u\|_2 = 1$)

$\|(x_0 - p)\|_2^2 + \lambda^2 \underbrace{\|u\|_2^2}_{1} + 2\lambda u^T (x_0 - p)$

at minimum $\nabla_\lambda(\cdot) = 0$

$\Leftrightarrow 2\lambda^* + 2u^T(x_0 - p)$

$\Leftrightarrow \lambda^* = u^T(p - x_0)$

$\therefore p^* = x_0 + u^T(p - x_0) u$ ☺

$a^T z_0 = b$

- Euclidian projection of a point (p) onto a hyperplane $H = \{z \in \mathbb{R}^n : a^T z = b\} = z_0 + \underbrace{\{z \in \mathbb{R}^n : a^T z = 0\}}_{\mathcal{N}(a^T)} = z_0 + \mathcal{N}(a^T)$

this is an affine set

$a \neq 0$ normal direction of the hyperplane

$\mathcal{N}(a^T)$ {which is a subspace}

$\forall_{z_1, z_2 \in H} \; \underbrace{a^T z_1 = b \wedge a^T z_2 = b}$

$\rightarrow a^T(z_1 - z_2) = 0 \Leftrightarrow a \perp (z_1 - z_2)$

$q \in H \Leftrightarrow a^T q = b$

$$p^* = \underset{q \in H}{\operatorname{argmin}} \|q - p\|_2^2$$

$$\underset{q \in H}{\forall} \|q - p\|_2^2 = \underset{\substack{q \\ a^T q = b}}{\forall} \|q - p\|_2^2 \quad L(q, \nu) = \|q - p\|_2^2 + \nu^T(a^T q - b) = \|q - p\|_2^2 + (a\nu)^T q - b^T \nu$$

$$g(\nu) = \underset{q}{\forall} \underbrace{\left(\|q - p\|_2^2 + (a\nu)^T q\right)}_{\nabla} - b^T \nu = \left\|-\frac{a\nu}{2} + \cancel{p} - \cancel{p}\right\|_2^2 + \underbrace{(a\nu)^T\left(-\frac{a\nu}{2} + p\right)} - b^T \nu = \frac{1}{4}(\nu^T a^T)(a\nu) - \frac{1}{2}\nu^T a^T a \nu + (p^T a)\nu - b^T \nu$$

$$= -\frac{1}{4}\nu^T a^T a \nu + (a^T p - b)^T \nu$$

$2(q - p) + a\nu = 0$

$\Leftrightarrow q - p = -\frac{a\nu}{2}$

$\Leftrightarrow p^* = q = -\frac{a\nu}{2} + p$

dual problem: $\underset{\nu}{R}\, g(\nu) = \underset{\nu}{R}(\cdot)$ ∇

$-\frac{1}{\cancel{2}} \cancel{2} \cdot (a^T a)\nu + a^T p - b = 0$

$\rightarrow \frac{1}{2}(a^T a)\nu = (a^T p - b)$

$\rightarrow \nu = 2\frac{a^T p - b}{\|a\|_2^2}$

$\therefore p^* = -\frac{a}{\cancel{2}} \cancel{2} \frac{(a^T p - b)}{a^T a} + p$

$= -a(a^T a)^{-1}(a^T p - b) + p$

$$\boxed{\therefore p^* = p - \frac{(a^T p - b)}{\|a\|_2^2} a}$$ Ans: [eq: Eucledian_proj_on_hyperplane]

By projection theorem:

$(p^* - p) \perp$ subspace_of_the_affine_set $= \mathcal{N}(a^T)$ $\{\because H = z^0 + \mathcal{N}(a^T)\}$

$\overset{\text{le}}{=} (p^* - p) \perp \mathcal{N}(a^T)$

$\overset{\text{le}}{=} \forall_{\eta \in \mathcal{N}(a^T)} \; (p^* - p)^T \eta = 0$

$\overset{\text{le}}{=} \forall_\eta \left(\underbrace{a^T \eta = 0}_{\text{given}} \Rightarrow \underbrace{(p^* - p)^T \eta = 0}_{\text{goal}}\right)$ Proof: $(p^* - p)^T \eta = -\frac{(a^T p - b)}{\|a\|_2^2} \underbrace{a^T \eta}_{=0} = 0$ ☺

2.3.2.3. Projection on a vector span:

$S = \operatorname{span}(x^{(1)}, \ldots, x^{(d)}) = \left\{\sum_{i=1}^{d} \alpha_i x^{(i)} \mid \alpha \in \mathbb{R}^d\right\}$ {subspace}

By projection theorem: (point_to_be_projected − projected_point) ⊥ (subspace_to_be_projected_on)

(p = point_to_be_projected, p^* = projected_point, S = subspace_to_be_projected_on)

$(p - p^*) \perp S$

$\overset{\text{le}}{=} \forall_{\omega \in S} \; (p - p^*)^T \omega = 0$

$\overset{\text{le}}{=} \forall_\omega \left(\left(\exists_\alpha \; \omega = \sum_{i=1}^{d} \alpha_i x^{(i)}\right) \rightarrow (p - p^*)^T \omega = 0\right)$

$\Rightarrow \forall_\omega \left(\underbrace{\exists_\alpha \; \omega = \sum_{i=1}^{d} \alpha_i x^{(i)}}_{\text{true as } S \neq \phi} \wedge (p - p^*)^T \omega = 0\right) \; [\because \exists_x (P(x) \wedge P(x) \Rightarrow Q(x)) \Rightarrow \exists_z (P(z) \wedge Q(z))]$

আমার equation ২ forall quantifier, so এখানে apply করা যাব.

$\Rightarrow \forall_\alpha \left((p - p^*)^T \sum_{i=1}^{d} \alpha_i x^{(i)} = 0\right) \overset{\text{le}}{=} \exists_\beta \forall_\alpha \left(p - \sum_{j=1}^{d} \beta_j x^{(j)}\right)^T \sum_{i=1}^{d} \alpha_i x^{(i)} = 0 \overset{\text{le}}{=} \exists_\beta \forall_\alpha \sum_{i=1}^{d} \alpha_i \left(p - \sum_{j=1}^{d} \beta_j x^{(j)}\right)^T x^{(i)} = 0 \overset{\text{le}}{=} \exists_\beta \forall_{i \in \{1, \ldots, d\}} \left(p - \sum_{j=1}^{d} \beta_j x^{(j)}\right)^T x^{(i)} = 0 \overset{\text{le}}{=} \exists_\beta \boxed{\forall_{i \in \{1, \ldots, d\}} \; p^T x^{(i)} = \sum_{j=1}^{d} \beta_j \left(x^{(j)T} x^{(i)}\right)}$ (eq: 2.8)

আবার $p^* \in S \Leftrightarrow \exists_\beta \; p^* = \sum_{j=1}^{d} \beta_j x^{(j)}$

যেহেতু
$\omega \rightarrow \alpha$ mapping one-to-one
ω এর পরিবর্তে α কে forall quantifier
হিসেবে ব্যবহার করা যাবে.

যেকোনো p^*, S এর কোন member select করা হলো তার উপর নির্ভর করবে না.

যখন α এর মান যাইহোক হবে যখন individually প্রতিটি term 0 হবে

Solving this linear equation set we can find out easily β hence $p^* = \sum_{j=1}^{d} \beta_j x^{(j)}$ ☺

*** Projection onto the span of orthonormal vectors:**

(eq: 2.8) থেকে:

$$\forall_{i \in \{1, \ldots, d\}} \; p^T x^{(i)} = \sum_{j=1}^{d} \beta_j \underbrace{x^{(j)T} x^{(i)}}_{\substack{1, \text{ if } i = j \\ 0, \text{ else}}} = \beta_i \|x^{(i)}\|_2^2 \rightarrow \beta_i = \frac{p^T x^{(i)}}{\|x^{(i)}\|_2^2}$$

$$\therefore p^* = \sum_{i=1}^{d} \beta_i x^{(i)} = \sum \frac{(p^T x^{(i)})}{\|x^{(i)}\|_2^2} x^{(i)}$$ ☺

* Projection on affine set $Az=b$: Projection on Az=b

↳ : fat, full rank

$$\mathcal{X}=\{z \mid Az=b\}$$

$$\therefore [x]_{\mathcal{X}} = \underset{z\in\mathcal{X}}{\operatorname{argmin}} \tfrac{1}{2}\|z-x\|_2^2 \xrightarrow{\text{opt problem}} \min \tfrac{1}{2}\|z-x\|_2^2 \ \text{s.t.}\ Az=b \xrightarrow{L(\square)} L(z,\nu)=\tfrac{1}{2}\|z-x\|_2^2+\nu^T(b-Az)$$

$$g(\nu)=\min_z \tfrac{1}{2}\|z-x\|_2^2+\underbrace{\nu^T(b-Az)}_{\nu^Tb-(A^T\nu)^Tz} \xrightarrow{\nabla(\square)=0} (z-x)-A^T\nu=0 \rightarrow z^*(\nu)=x+A^T\nu$$

$$=\left[\tfrac{1}{2}\|z-x\|_2^2+\nu^Tb-(A^T\nu)^Tz\right]_{z=x+A^T\nu}$$

$$=\tfrac{1}{2}\|A^T\nu\|_2^2+\nu^Tb-\underbrace{(A^T\nu)^T(A^T\nu+x)}_{\|A^T\nu\|_2^2+\underbrace{\nu^TAx}_{(Ax)^T\nu}}=\tfrac{1}{2}\|A^T\nu\|_2^2+b^T\nu-\|A^T\nu\|_2^2-(Ax)^T\nu=-\tfrac{1}{2}\|A^T\nu\|_2^2+(b-Ax)^T\nu$$

⇓ dual problem

$$\max\ -\tfrac{1}{2}\|A^T\nu\|_2^2+(b-Ax)^T\nu$$

$$=-\max\underbrace{\left(\tfrac{1}{2}\|A^T\nu\|_2^2-(b-Ax)^T\nu\right)}$$

⇓ $\nabla_\nu(\square)=0$

$$(A^T)^T(A^T\nu)-(b-Ax)=0$$

$$\rightarrow (AA^T)\nu=(b-Ax)$$

$$\rightarrow \nu^*=(AA^T)^{-1}(b-Ax)$$

$$z^*=x+A^T\nu^*$$

$$=x+A^T(AA^T)^{-1}(b-Ax)$$

$$=x+A^T(AA^T)^{-1}b-A^T(AA^T)^{-1}Ax$$

$$=(I-A^T(AA^T)^{-1}A)x+A^T(AA^T)^{-1}b$$

$$\therefore \Pi_{\{\square\mid A\square=b\}}(x)=[x]_{\{\square:\,A\square=b\}}=(I-A^T(AA^T)^{-1}A)x+A^T(AA^T)^{-1}b$$

\# Projection on A z = b