Adjoint of linear maps

Let $A: U \xrightarrow{\text{linear}} V$ be a linear map

Def $A^*: V^* \xrightarrow{\text{linear}} U^*$ called the adjoint of A is defined by

$\langle A^*\lambda | u \rangle = \langle \lambda | Au \rangle \quad \forall \lambda \in V^*$ and $u \in U$.

If $\vec{e}_1, \cdots, \vec{e}_n$ for U and $\vec{f}_1, \cdots, \vec{f}_m$ for V

$\updownarrow$ dual basis $\qquad\qquad$ $\updownarrow$ dual basis

$dx^1, \cdots, dx^n$ for U^* $\qquad$ $dy^1, \cdots, dy^m$ for V^*

$$A = a^i_j \vec{f}_i dx^j$$

- $\vec{u} = u^k \vec{e}_k$ then $A\vec{u} = a^i_j \vec{f}_i \langle dx^j | u^k \vec{e}_k \rangle$

$$= a^i_j \vec{f}_i \delta^j_k u^k = a^i_j u^j \vec{f}_i$$

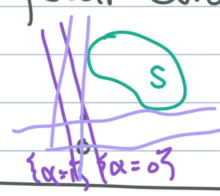
- $\lambda = \lambda_l dy^l$. Then $A^*\lambda = a^i_j \lambda_i$.

Def If $W \subset U$ is a subspace, let the annihilator space be

$$W^\circ \subset U^* \qquad W^\circ := \{\lambda \mid \langle \lambda | \vec{w} \rangle = 0 \ \forall \vec{w} \in W.\}$$

Generalization If $S \subset U$ is a subset, let the polar cone $S^\circ \subset U^*$ be

$$S^\circ := \{\lambda \mid \langle \lambda | \vec{s} \rangle \leq 0 \ \forall \vec{s} \in S\}$$



elements of S° are covectors whose positive side don't intersect with S.

Prop S° is always a convex cone.

Four fundamental subspaces.

Let $A: U \xrightarrow{\text{linear}} V$ be a linear map. Its 4 fundamental spaces are

$$\ker(A) = \{\vec{u} \in U \mid A\vec{u} = 0\} \subset U$$
$$\operatorname{im}(A) = \{A\vec{u} \mid \vec{u} \in U\} \subset V$$
$$\ker(A^*) = \{\lambda \in V^* \mid A^*\lambda = 0\} \subset V^*$$
$$\operatorname{im}(A^*) = \{A^*\lambda \mid \lambda \in V^*\} \subset U^*.$$

Thm $\ker(A)^\circ = \operatorname{im}(A^*) \qquad \operatorname{im}(A)^\circ = \ker(A^*)$

$\ker(A^*)^\circ = \operatorname{im}(A) \qquad \operatorname{im}(A^*)^\circ = \ker(A)$

Optimization problems

Optimality condition for optimization

① Unconstrained problem

Let M be a domain without boundary (e.g. $M=\mathbb{R}^n$)

Let $E: M \to \mathbb{R}$, smooth function.

Consider $\boxed{\underset{x\in M}{\text{minimize}}\; E(x)}$

Optimality condition: If $x_0 \in M$ is the minimizer, then

$$\boxed{dE|_{x_0}[\![\mathring{x}]\!]=0 \quad \text{for all } \mathring{x}\in T_{x_0}M.}$$

Equivalently,

$\boxed{dE|_{x_0}=0}$ or $\boxed{\text{grad} E|_{x_0}=0}$

using any inner product structure

② Equality constrain

Let M be a domain without boundary. (e.g. $M=\mathbb{R}^n$)

Let $S\subset M$ be a surface without boundary (e.g. $S=\{x\in M \mid g(x)=0\}$ where $g: M\to\mathbb{R}^m$)

Consider $\boxed{\underset{x\in S}{\text{minimize}}\; E(x)}$

By ① we know the optimality condition for $x_0\in S$ is

$$dE|_{x_0}[\![\mathring{x}]\!]=0 \quad \forall\; \mathring{x}\in T_{x_0}S$$

or just $dE|_{x_0} \in T_{x_0}S^{\circ}$

annihilator

Concretly: Let $g: M \to Y$ for some vector space Y, which defines the constraint $S = \{x \in M \mid g(x) = 0_Y\}$.

$$\boxed{\underset{x \in M}{\text{minimize}}\ E(x) \quad \text{subject to } g(x) = 0}$$

What is T_xS? $T_xS = \{\mathring{x} \in T_xM \mid dg[\![\mathring{x}]\!] = 0\} = \ker(dg|_x)$

What is $(T_xS)^\circ$? $\ker(dg|_x)^\circ = \operatorname{im}(dg|_x^*)$

$= \{dg|_x^*[\![\lambda]\!] \mid \lambda \in Y^*\}$

Optimality: $\exists \lambda \in Y^*$ (called Lagrange multiplier)

so that $dE|_{x_0} = dg|_{x_0}^* \lambda$.

In coordinate:

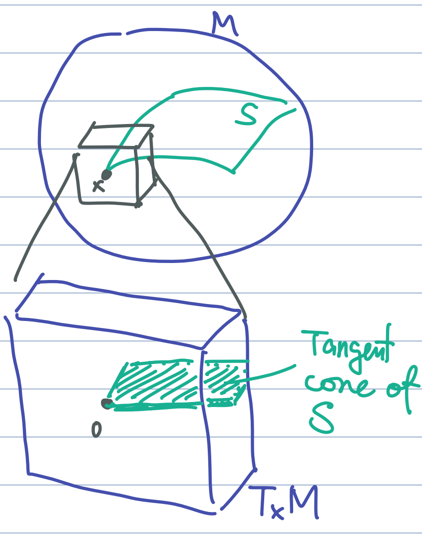
$$\left[\frac{\partial E}{\partial x^1} \cdots \frac{\partial E}{\partial x^n}\right]\Bigg|_{x_0} = [\lambda_1 \cdots \lambda_m] \begin{bmatrix} \frac{\partial g^1}{\partial x^1} & \cdots & \frac{\partial g^1}{\partial x^n} \\ & \vdots & \\ \frac{\partial g^m}{\partial x^1} & \cdots & \frac{\partial g^m}{\partial x^n} \end{bmatrix}\Bigg|_{x_0}$$

or $$\frac{\partial E}{\partial x^i} = \lambda_\alpha \frac{\partial g^\alpha}{\partial x^i}.$$

③ Inequality constraints

Let $S \subset M$ be a surface with boundary.

(e.g. $S = \{x \in X \mid h_i(x) \le 0 \quad i = 1, \cdots, l\}$)

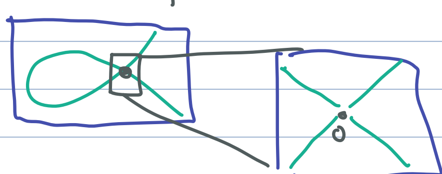


For each $x \in S$ let the tangent cone of S at x

$$C_x := \left\{\mathring{x} \in T_xM \;\middle|\; \mathring{x} = \frac{d}{dt}\gamma(t)\Big|_{t=0} \text{ for some } \gamma: [0,1] \to S\right\}$$

C_x could be a subspace or the entire space.

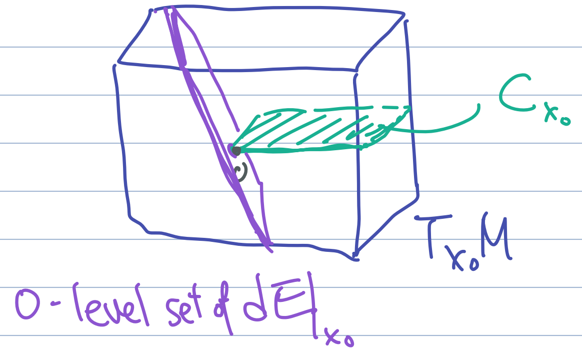
C_x could be non-convex.



Optimality condition for $\boxed{\min_{x\in S} E(x)}$ is

$$\boxed{-dE|_{x_0} \in C^{\circ}_{x_0}}$$

Recall $C^{\circ}_{x_0} = \{ \lambda \in T^*_{x_0}M \mid \langle \lambda | \vec{c}\rangle \le 0 \ \forall \vec{c}\in C_{x_0} \}$



A general optimization problem written in coordinates is

$$\begin{cases} \min_{x\in M=\mathbb{R}^n} E(x) & & (E: M\to\mathbb{R}) \\ g^i(x) = 0 & i=1,\cdots,m & (g: M \to Y = \mathbb{R}^m) \\ h^j(x) \le 0 & j = 1,\cdots,l & (h: M\to Z = \mathbb{R}^l) \end{cases}$$

<u>Karush-Kuhn-Tucker conditions</u> (KKT) (1939, 1951)

$\exists$ Lagrange multipliers $\lambda = [\lambda_1,\cdots,\lambda_m] \in Y^*$

$\mu = [\mu_1,\cdots,\mu_l] \in Z^*$ so that

$$\begin{cases} dE|_{x_0} + (dg|_{x_0})^*\lambda + (dh|_{x_0})^*\mu = 0 \\ \mu_1,\cdots,\mu_l \ge 0 \\ \langle \mu | h(x_0)\rangle = 0 \quad (h(x_0)\in Z \text{ is contained in the 0-level set of covector } \mu) \\ h(x_0) \le 0, \ g_{x_0} = 0. \end{cases}$$