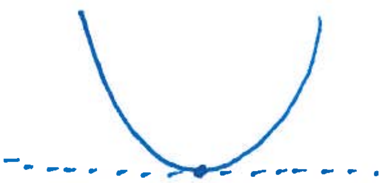
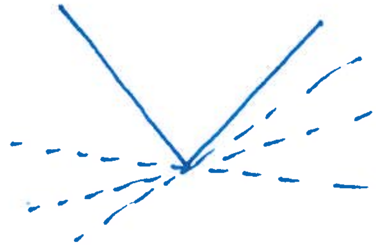
Subgradients (for convex functions)

Derivative/gradient notion for nonsmooth convex functions





v is <u>subgradient</u> of $f: \mathbb{R}^n \to \mathbb{R}$ at y if $f(x) - f(y) \geq \langle v, x-y \rangle \ \forall x$

The subdifferential at y, $\partial f(y)$, is set of all subgradients at y. (set valued notion of derivative)

If f smooth $\partial f(y) = \{ \nabla f(y) \}$

If $f(x) = |x|$, $\partial f(0) = [-1, 1]$

If $f(x) = \|x\|_1$ for $x \in \mathbb{R}^n$, $\partial f(0) = \{ x \mid \|x\|_\infty \leq 1 \}$

$$\partial f(y) = \left\{ x \;\middle|\; \begin{array}{ll} x_i = \mathrm{sign}(y_i) & \text{if } y_i \neq 0 \\ |x_i| \leq 1 & \text{if } y_i = 0 \end{array} \right.$$

Derivation of Dual Problem to basis pursuit

$\min \|x\|_1$ st $Ax = b$

$\mathcal{L}(x, \lambda) = \|x\|_1 - \langle \lambda, Ax - b \rangle$

$$g(\lambda) = \inf_x \mathcal{L}(x, \lambda) = \inf \|x\|_1 - \langle A^t\lambda, x\rangle + \langle \lambda, b\rangle$$
$$= \langle \lambda, b\rangle + \inf_x \|x\|_1 - \langle A^t\lambda, x\rangle$$

Study $\inf_x \|x\|_1 - \langle A^t\lambda, x\rangle$. when is it $-\infty$? when finite.

If $\|A^t\lambda\|_\infty > 1$, $\inf = -\infty$. Just choose x that selects largest index of $A^t\lambda$.

What does it take to minimize $\|x\|_1 - \langle A^t\lambda, x\rangle = f(x)$

$0 \in \partial f(x) \Rightarrow 0 \in \partial\|x\|_1 - A^t\lambda. \Rightarrow A^t\lambda \in \partial\|x\|_1(x)$

So $(A^t\lambda)_i = \begin{cases} \operatorname{sign} x_i & \text{if } x_i \neq 0 \\ \in [-1, 1] & \text{if } x_i = 0 \end{cases}$ possible if $\|A^t\lambda\|_\infty \leq 1$

~~Choose $x = \operatorname{sign}(A^t\lambda)$. So $\|x\|_1 - \langle A^t\lambda, x\rangle = \|x\|_1 - \langle A^t\lambda, \operatorname{sign} A^t\lambda\rangle = \|\operatorname{sign} A^t\lambda\|_1$~~

$\|x\|_1 - \langle A^t\lambda, x\rangle = \|x\|_1 - \langle \operatorname{sign} x, x\rangle = \|x\|_1 - \|x\|_1 = 0.$

So $\inf_x \|x\|_1 - \langle A^t\lambda, x\rangle = \begin{cases} 0 & \text{if } \|A^t\lambda\|_\infty \leq 1 \\ -\infty & \text{otherwise} \end{cases}$

Dual problem

$$\sup_\lambda g(\lambda) = \sup \langle \lambda, b\rangle \text{ st } \|A^t\lambda\|_\infty \leq 1$$

By strong duality of LPs, $\sup_\lambda g(\lambda) = \inf \|x\|_1$ st $Ax = b$

Activity:

What is $\partial f(x)$ for $f(x) = \|x\|_2$?

Optimality by dual certification

$\min \|X\|_1$ st $Ax = b$ $(*)$

$\mathcal{L}(X,\lambda) = \|X\|_1 - \langle \lambda, Ax - b \rangle$

Optimality condition at x_0

$0 \in \partial\|\cdot\|_1(x_0) - A^t\lambda \Rightarrow (A^t\lambda)_S = \text{sign}(X_{0,S})$ on $S = \text{support } X_0$

$\|A^t_{S^c}\lambda\|_\infty \leq 1$

Claim: Let $S = \text{supp}(X_0)$. If $\exists \lambda$ st $A^t_S\lambda = \text{Sign } X_{0,S}$ on

$\|A^t_{S^c}\lambda\|_\infty \leq 1$

then X_0 is a minimizer to $(*)$

Proof: By duality theory, dual optimum is bdd from above by primal optimum.

By assumptn $\exists \lambda$ st $\|A^t\lambda\|_\infty \leq 1$. So λ dual feasible

$\langle \lambda, b \rangle = \langle \lambda, AX_0 \rangle = \langle A^t\lambda, X_0 \rangle = \|X_0\|_1.$

So $\|X_0\|_1$ is optimal primal obj. So X_0 is a minimizer.

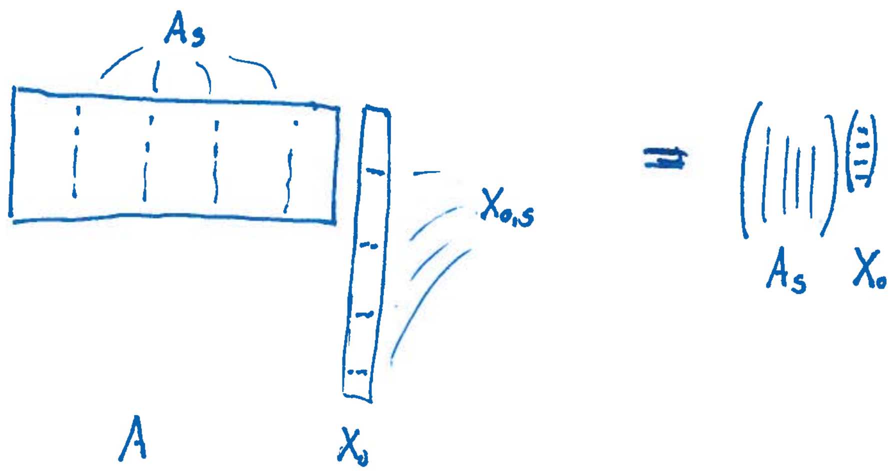
No guarantee of uniqueness.

Theorem: Let $S = \text{support}(X_0)$. Suppose A_S has full rank.

Suppose $\exists \lambda$ st $A_S^t \lambda = \text{sign}(X_{0,S})$

$$\|A_{S^c}^t \lambda\|_\infty < 1$$

Then X_0 is unique soln to $\min \|X\|_1$ st $AX = AX_0$.



Proof: Let $Y = A^t \lambda$.

Consider $\|X_0 + h\|_1 \le \|X_0\|_1$ st $Ah = 0$

$$0 = \langle Ah, \lambda \rangle = \langle h, A^t \lambda \rangle = \langle h_S, \text{sign}\, X_{0,S} \rangle + \langle h_{S^c}, Y_{S^c} \rangle$$

$$\Rightarrow \quad \langle h_S, \text{sign}\, X_{0,S} \rangle = -\langle h_{S^c}, Y_{S^c} \rangle \qquad (*)$$

As $\|X_0 + h\|_1 \le \|X_0\|_1$,

$\|X_0 + h_S\|_1 + \|h_{S^c}\|_1 \le \|X_0\|_1$

$\|X_0\|_1 + \langle \text{sign}\, X_{0,S}, h_S \rangle_1 + \|h_{S^c}\|_1 \le \|X_0\|_1$

$$\langle \text{sign}\, X_{0,S}, h_S \rangle + \|h_{S^c}\|_1 \le 0$$

$$\Rightarrow \quad \|h_{S^c}\|_1 \le \langle h_{S^c}, Y_{S^c} \rangle \le \|h_{S^c}\|_1 \|Y_{S^c}\|_\infty \qquad \text{by } (*)$$

$$\Rightarrow \quad \|h_{S^c}\|_1 \left[1 - \|Y_{S^c}\|_\infty \right] \le 0 \Rightarrow \|h_{S^c}\|_1 = 0 \Rightarrow h_{S^c} = 0.$$

Hence, soln lives on S. As A_S has full rank, $A_S h_S = 0 \Rightarrow h_S = 0$. So $h = 0$ ■

How to find a dual certificate

Recovery guarantee if $\exists\, \lambda$ st $\begin{cases} A_S^t \lambda = sgn(X_{0,S}) \\ |A_{S^c}^t \lambda| < 1 \end{cases}$

How to build such a λ?

Choose smallest λ (in ℓ_2 sense) that fits constraints.

$$\min \|\lambda\|_2 \quad st \quad A_S^t \lambda = sgn(X_{0,S})$$ underdetermined system

$$\lambda = A_S (A_S^t A_S)^{-1} sgn(X_{0,S})$$

Suffice if $\| A_{S^c}^t A_S (A_S^t A_S)^{-1} sgn\, X_{0,S} \|_\infty < 1$.

Thm: Let $M = \sup_{i \neq j} a_i \cdot a_j$ w/ $\|a_i\| = 1$ (coherence)

If $\|X_0\|_0 \leq \frac{1}{2}(1 + \frac{1}{M})$ then $\| A_{S^c}^t A_S (A_S^t A_S)^{-1} sgn\, X_{0,S} \|_\infty < 1$.

So X_0 is unique soln to $\min \|X\|_1$ st $Ax = Ax_0$.