

$$\min_{x \in C} g(x) + h(x)$$

$$\min_x \underbrace{g(x)}_{\text{+}} + \underbrace{h(x) + I_C(x)}_{\text{+}}$$

↑  
need prox of this

$$\text{prox}_t(x) = \arg \min_z \frac{1}{2t} \|x - z\|_2^2 + h(z) + I_C(z)$$

$$= \arg \min_{z \in C} \frac{1}{2t} \|x - z\|_2^2 + h(z)$$

e.g.  $C = \{x : Ax \leq b\}$

$$h(x) = \|x\|_1$$

prox of  $h$  is easy here; but not constrained prox

$$\min f(x)$$

s.t.  $\underbrace{\|x\|_1 \leq R}$

$$x^+ = P_{\{z : \|z\|_1 \leq R\}}(x - t \nabla g(x))$$

$$\min_x f(x) + \|x\|_1 \iff \min_{x, z} f(x) + 1^T z$$

$-z \leq x \leq z$

$$\min f(x)$$

s.t.  $h_i(x) \leq 0$

approximate  
with

$$\min_x f(x) \leftarrow -\frac{1}{t} \sum_i \log(-h_i(x))$$

check if

$$f(x + tv) > f(x) + \alpha t \nabla f(x)^T v$$

passes, shrink  $t$

else, stop.

---

E.g. compare for this problem

$$\min_{\beta} \sum_{i=1}^m -y_i x_i^T \beta + \log(1 + e^{x_i^T \beta}) + \lambda \|\beta\|_1$$

all the methods that are applicable.

strengths & weaknesses

$$(f) \quad x_1 + x_2 = -2x_1 + x_3 = tb$$

$$\log(\underline{\quad} + \underline{\quad}) = \log(2e^{tb}) \\ =$$

$\log(\sum e^{x_i})$  log-sum-exp

Taylor Series with Remainder

$$f(y) = f(x) + f'(x)(y-x) + \frac{f''(x)}{2}(y-x)^2 \\ + \underline{\text{Rem}}$$

e.g.

$$f(y) = f(x) + f'(x)(y-x) + \frac{f''(z)}{2}(y-x)^2$$

for some  $z$  between  $x$  and  $y$ .

~~WRECKED~~

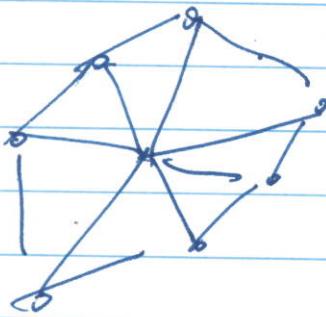
$$\min_{\mathbf{X}} \|\mathbf{X} - \mathbf{R}\|_F^2 \quad \text{s.t. } \text{rank}(\mathbf{R}) = k$$

$$\Leftrightarrow \min_{\mathbf{Z}} \|\mathbf{X} - \mathbf{XZ}\|_F^2 \quad \text{s.t. } \text{rank}(\mathbf{Z}) = k \\ \mathbf{Z} \text{ is a proj.}$$

$$\|\mathbf{X} - \mathbf{R}\|_F^2 = \|\mathbf{X} - \mathbf{R}_1\|_F^2 + \|\mathbf{R}_2\|_F^2$$

$$\mathbf{R} = \mathbf{R}_1 + \mathbf{R}_2$$

$\uparrow$   
aligned with col space of  $\mathbf{X}$



$$\max_{\mathbf{Z}} \langle \mathbf{x}, \mathbf{z} \rangle$$

$$\text{s.t. } \text{rank}(\mathbf{Z}) = k$$

$\mathbf{Z}$  proj.

$$\Leftrightarrow \max_{\mathbf{Z}} \langle \mathbf{x}, \mathbf{z} \rangle$$

$$\text{s.t. } \text{tr}(\mathbf{Z}) = k$$

$$0 \leq \mathbf{Z} \leq \mathbf{I}$$

Prox operator. How to derive?

1. Does it decompose? Across coordinates? Block?
2. Argue based on first principles?
3. Take subgradients, set equal to 0.

$$\min_{\mathbf{Z}} \frac{1}{2} \|\mathbf{x} - \mathbf{z}\|_2^2 + t h(\mathbf{z})$$

$$\mathbf{z} - \mathbf{x} + t \nabla h(\mathbf{z}) = 0$$

e.g.  $\ell_1$  norm

$$\min_{\mathbf{Z}} \frac{1}{2} \|\mathbf{x} - \mathbf{z}\|_2^2 + t \|\mathbf{z}\|_1$$

$$\Leftrightarrow \min_{z_i} \frac{1}{2} (x_i - z_i)^2 + t |z_i|$$

$i=1, \dots, n$

e.g.  $\ell_2$  norm

$$\min_{\mathbf{Z}} \frac{1}{2} \|\mathbf{x} - \mathbf{z}\|_2^2 + t \|\mathbf{z}\|_2$$

$$\mathbf{z} = c\mathbf{x}$$

