#3.1

$\underline{\underline{A}} \in \mathbb{R}^{m\times n}$, $\underline{b} \in \mathbb{R}^n$, $\underline{\underline{L}} \in \mathbb{R}^{p\times n}$, $\lambda \in \mathbb{R}_{++}$

Regularized LS:

$$\min_{\underline{x} \in \mathbb{R}^n} \|\underline{\underline{A}}\underline{x} - \underline{b}\|_2^2 + \lambda \|\underline{\underline{L}}\underline{x}\|_2^2 = \min_{\underline{x}} f(\underline{x})$$

Q: show that this problem has a unique sol'n iff $Null(\underline{\underline{A}}) \cap Null(\underline{\underline{L}}) = \{\underline{0}\}$.

$$f(\underline{x}) = \overset{(\underline{\underline{A}}\underline{x} - \underline{b})^T}{\cancel{(\underline{b} - \underline{\underline{A}}\underline{x})^T}} (\underline{\underline{A}}\underline{x} - \underline{b}) + \lambda \underline{x}^T \underline{\underline{L}}^T \underline{\underline{L}} \underline{x}$$

$$\nabla^2 f(\underline{x}) = 2\underline{\underline{A}}^T\underline{\underline{A}} + 2\overset{\lambda}{\vee}\underline{\underline{L}}^T\underline{\underline{L}} = 2(\underline{\underline{A}}^T\underline{\underline{A}} + \overset{\lambda}{\vee}\underline{\underline{L}}^T\underline{\underline{L}}).$$

$\nabla^2 f \succ \underline{0} \iff$ reg. LS has a unique solution

When is $\nabla^2 f \succ \underline{0}$?

I.e. when is $\underline{y}^T \nabla^2 f \underline{y} > 0 \ \forall \ \underline{y} \neq \underline{0}$?

when is $\underline{y}^T \underline{\underline{A}}^T \underline{\underline{A}} \underline{y} + \lambda \underline{y}^T \underline{\underline{L}}^T \underline{\underline{L}} \underline{y} > 0$ if $\underline{y} \neq \underline{0}$?

when is $\|\underline{\underline{A}}\underline{y}\|_2^2 + \lambda \|\underline{\underline{L}}\underline{y}\|_2^2 > 0$ if $\underline{y} \neq \underline{0}$?

$\|\underline{\underline{A}}\underline{y}\|_2 = 0$ iff $\underline{\underline{A}}\underline{y} = \underline{0}$

$\|\underline{\underline{L}}\underline{y}\|_2 = 0$ iff $\underline{\underline{L}}\underline{y} = \underline{0}$

If there does not exist $\underline{y} \neq \underline{0}$ s.t. both $\underline{\underline{A}}\,\underline{y} = \underline{0}$ and $\underline{\underline{L}}\,\underline{y} = \underline{0}$, then $\nabla^2 f \succ \underline{\underline{0}}$.

$\Updownarrow$

$\text{Null}(\underline{\underline{A}}) \cap \text{Null}(\underline{\underline{L}}) = \{\underline{0}\}$

$\Updownarrow$

regularized LS problem has a unique solution.

Hw #3, P2

$$\mathcal{L}(p) = \binom{100}{H} p^H (1-p)^T, \quad H + T = 100$$

$$p \in [0, 1].$$

For global optimization of continuous functions, the max occurs either

✓ (i) on the interior, $p \in (0, 1)$

✗ (ii) on the boundary, $p = 0, 1$.

$(\mathcal{L}(0) = \mathcal{L}(1) = 0)$

$$\max_{p \in (0,1)} \mathcal{L}(p) \iff \max_{p \in (0,1)} \log \mathcal{L}(p)$$

since $\log(\cdot)$ is strictly monotonic.

$$\underset{\substack{\parallel \\ f(p)}}{\log \mathcal{L}(p)} = \log \binom{100}{H} + H \log p + T \log(1-p)$$

$$f'(p) = \frac{H}{p} - \frac{T}{1-p}$$

stationary pts: $f'(p) = 0$

$$(1-p)H = Tp$$

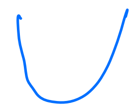
$$p = \frac{H}{H+T} = \frac{H}{100}$$

$$f''(p) = \frac{-H}{p^2} - \frac{T}{(1-p)^2} < 0 \quad \text{if } p \in (0,1)$$

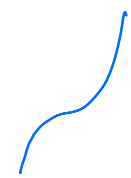
So: $p = \frac{H}{100}$ is a local max (and global one since no other stationary pts)

<u>Coercivity</u>: $\left(\lim_{\|\mathbf{x}\| \to \infty} f(\mathbf{x}) = +\infty \right)$

<u>Ex</u>: $f(x) = x^2$ is coercive



$f(x) = x|x|$ is <u>not</u> coercive



$f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ and $\mathbf{A} \succ \mathbf{0}$, then f is coercive.

if $\mathbf{A} \prec \mathbf{0}$, then $-f$ is coercive

if $\mathbf{A}$ is indefinite, then it's not coercive.

<u>Ex</u>. $f(\mathbf{x}) = x_1^4 + x_2^4 = \|\mathbf{x}\|_4^4$

outside $B[\mathbf{0}, 1]$, then $x_1^4 + x_2^4 \geq x_1^2 + x_2^2 = \|\mathbf{x}\|_2^2$

$f(\underline{x}) = x_1^3 + x_2^3$. (not coercive)

$$\lim_{x_1 \to \infty} f(-x_1, 0) = (-x_1)^3 \to -\infty.$$

$$f(\underline{x}) = \left(\sum_{i=1}^{n} x_i\right)^2$$

$$\lim_{x_1 \to \infty} f(x_1, -x_1, 0, 0, 0 \ldots) = \lim_{x_1 \to \infty} (x_1 - x_1)^2$$

$$= 0.$$

Cauchy-Schwarz inequality: $\underline{x}, y \in \mathbb{R}^n$

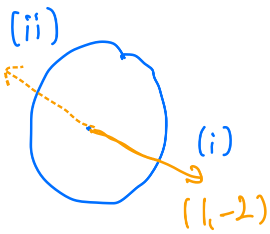
$$|\underline{x}^T y| = |\langle \underline{x}, y \rangle| \leq \|\underline{x}\|_2 \cdot \|y\|_2$$

$$\Updownarrow$$

$$-\|\underline{x}\|_2 \|y\|_2 \leq (\langle \underline{x}, y \rangle) \leq \|\underline{x}\|_2 \cdot \|y\|_2$$

with equality iff $\underline{x}$ and y are parallel.

__Ex__. optimize $f(x_1, x_2) = x_1^2 + x_2^2 + x_1 - 2x_2$ on $B[\underline{0}, 1]$



$$f(x_1, x_2) \le x_1^2 + x_2^2 + \left\| \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\|_2 \cdot \underbrace{\left\| \begin{pmatrix} 1 \\ -2 \end{pmatrix} \right\|_2}_{\sqrt{5}}$$

(i)

C-S.

(ii)

$$f(x_1, x_2) \ge x_1^2 + x_2^2 - \left\| \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\|_2 \cdot \left\| \begin{pmatrix} 1 \\ -2 \end{pmatrix} \right\|_2$$

In (i): equality when $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = k \begin{pmatrix} 1 \\ -2 \end{pmatrix}$, $k \in \mathbb{R}_{++}$

(ii) equality when $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = k \begin{pmatrix} -1 \\ 2 \end{pmatrix}$, $k \in \mathbb{R}_{++}$

Defining $r = \left\| \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\|_2$

Then along (i): $f(x_1, x_2) = r^2 + r\sqrt{5}$, $r \in [0, 1]$

since $\underline{x} \in B[\underline{0}, 1]$

$$\max_{\underline{x} \in B[\underline{0},1]} f(\underline{x}) = \max_{r \in [0,1]} r^2 + r\sqrt{5} \ \ldots$$

Similarly for (ii):

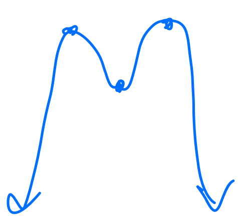
$$\min f(\underline{x}) = \min_{r \in [0,1]} r^2 - r\sqrt{5}$$

$$f(\underline{x}) \approx f(\underline{x}_0) + \underbrace{\nabla f(\underline{x}_0)^T (\underline{x} - \underline{x}_0)}_{\leq \|\nabla f(\underline{x}_0)\|_2 \cdot \|\underline{x} - \underline{x}_0\|_2}$$

$$\min_{x \in \mathbb{R}} f(x) \quad , \quad f \text{ is smooth}$$

Suppose f has a ~~single~~ stationary point that is a local min, and f at this stationary point is smaller than the value of f at all other stationary points.

Global min? ?

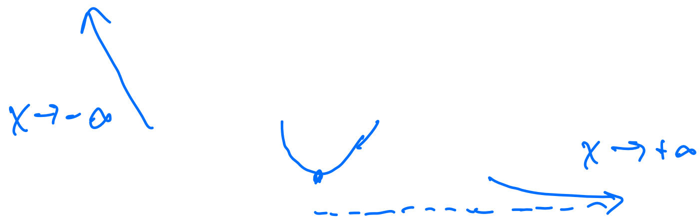


How to handle $\pm\infty$?

(i) coercitivity

(ii) "understand" how f behaves at ∞.

(iii) characterize $\nabla^2 f$



then min is "at"

$x = \infty$.