Lecture 6

Goal today: Basics of constrained optimization via constrained least squares.

- Quick review of conditions
- Intro to convexity and why it makes optimization easier
- Local mins are global mins for convex functions
- (pt) is both necessary and sufficient for continuous convex functions

We'll need a result from your HW to show least squares is convex.

- Solving least squares
- Constrained least squares
- Solving constrained least squares we probably won't quite get to today.

Logistics: Office hours
David: Fri 1:30-2:30 Lawson
Diana: Mon 12:30-12:30 DIAG

Unconstrained Convex Cond.

$f(x) \geq f(x^*) + \nabla f(x^*)^T(x - x^*)$

Necessary Cond for local min.

$g(x^*) = 0$ (no ∇ in HW)

$H(x^*) \succeq 0$ (convex)

Sufficient Cond.

$g(x^*) = 0$

$H(x^*) \succ 0$

then x^* is local min

Strong

Convexity

$f(x)$ is convex if for $\forall \theta \in [0,1]$ and $\forall x, y \in \mathbb{R}^n$

$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y)$

then f is convex.

Above: f must also be a convex set

Not Convex

If x^* is a local min

$\Rightarrow$ x^* is a global min of f too.

Proof: Suppose not. x^* is a local min

y^* is a global min $\neq x^*$

$f(x^*) > f(y^*)$

Look at $x_\theta = \theta x^* + (1-\theta) y^*$

$f(x_\theta) \leq \theta f(x^*) + (1-\theta) f(y^*)$

$< \theta f(x^*) + (1-\theta) f(x^*)$

$= f(x^*)$

$\Rightarrow f(x_\theta) < f(x^*)$

$\Rightarrow$ x^* is not a local min!

Result 2:

Let $f(x)$ be convex. Cont. diff.

$g(x_0) = 0$ then x_0 is a local (global) min

$\Leftrightarrow g(x_0) = 0$ is both sufficient for convexity.

Necessary and Sufficient for x_0 to be a local min (then $g = 0$ is also)

Proof:

Look at x_θ be a point where $g(x_0) = 0$ and x_0^* is not a local min.

Then not sufficient is exactly y such that

$f(x_0^*) \leq f(x_\theta)$

But $g(x_\theta) = 0$

Least Squares

$f(x) = \|Ax - b\|_2^2$

$= \|b - Ax\|_2^2$

$= \sum_{i=1}^{m} (b_i - a_i^T x)^2$

A has m rows, a_i^T is the ith row of A.

Solve LS $\Leftrightarrow$

find x to minimize $f(x)$.

$f(x)$ is convex. Will show $f(x)$ is convex.

How to solve LS

- Pseudo inverse: SVD
- Normal Equations
- Gram Matrix inverse
- QR
- Gradient Descent
- Conjugate Gradient
- LSQR
- CGLS
- LSMR
- Any method for solving linear eq

Show the convex

$g(x) = 0$

will show $f(x)$

$f(x) = \sum_{i=1}^{m} (a_i^T x - b_i)^2$

$A = \begin{bmatrix} a_1^T \\ \vdots \\ a_m^T \end{bmatrix}$ $b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$

$g(x) = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{bmatrix}$

$\frac{\partial f}{\partial x_k} = \frac{\partial}{\partial x_k} \sum_{i=1}^{m} (a_i^T x - b_i)^2$

$\frac{\partial}{\partial x_k} (a_i^T x - b_i)^2$

$2(a_i^T x_1 + b_i) a_{ik}$.

$f(x) = \theta Ax - b\|_2^2$

$= (Ax - b)^T (Ax - b)$

$= x^T A^T A x - 2x^T A^T b$

$g(x) = 2A^T A x - 2A^T b$

$g(x) = 0 \Leftrightarrow$

$A^T A x = A^T b$

Solving LS

find x s.t.

$g(x) = 0 \Leftrightarrow$

$2A^T A x - 2A^T b = 0$

$\Leftrightarrow$

Solve $A^T A x = A^T b$

Called the normal eq.

If A is full rank, $m \geq n$, then has a unique soln.

Quick quiz

What is g that solves

$\min \left\| \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} g - \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\|_2$

$\begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} g = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \Leftrightarrow \min \|Eq - b\|_2^2$

$A + E$

$A^T A g = A^T b \Rightarrow \frac{1}{2} \frac{1}{2} g = \frac{1}{2} \frac{1}{2}$

$\Leftrightarrow \frac{1}{4} g = \frac{1}{4}$

$g = \frac{1}{2}, \frac{1}{2}$

We know we want to consider Constrain least sq.

$\min \|Ax - b\|_2^2$

s.t. $Cx = d$

Look that Subject to

A: $m \times n$

C: $p \times n$.

Eg. $\min \|Ax - b\|$

s.t. $\sum x_i = 1$.

Q. How do we know find x s.t. $Cx = d$?

Find a feasible point $\Rightarrow$ A feasible point is any point that satisfies the Constraint.

Q2. How to solve CLS ?

3. Convex for checkers

- full rank C.

C is full $(p \times n)$ $\square \to \mathcal{N}$

C is square $(p \times n)$ $\square \to 1$

C is wide $(p < n)$ $\square \to 6$

What is the hardest (that we can still make)?

If C is square, $p = n$, and full rank, $Cx = d$ has exactly one soln.

C is wide $(p < n)$

$Cx = d$

has infinite solns

$\Rightarrow$ Underdetermined

If C is full rank, then Cx is a surjective

y s.t. $\exists x$ s.t. $Cx = y$.

Null$(C) = \{y \mid Cy = 0\}$

Note: if

$Cx_1 = d$

$Cx_2 = d$

for some x_1, x_2 then

$C(x_1 - x_2) = 0$

for all $y \in$ Null(C).

Min $\|b - Ax\|_2^2$

s.t. $Cx = d$

$\Leftrightarrow$

Min $\|b - Ax\|$

s.t. $Cx = d$

$\Leftrightarrow$

$C^T = V \Sigma$

C is $1 \times n$.

If C is wide, then Cx is a surjective

y s.t. $\exists x$ s.t. $Cx = y$.

Null$(C) = \{y \mid Cy = 0\}$

Null(C) is a subspace

$\dim$ Null$(C) = n - p$

Null(C) is $n - p$ dim.

Let Z be basis for null(C).

$\Rightarrow y \in$ Null$(C) \Leftrightarrow y = Zt$

Null space is a vector space

vector space has a basis. (dim. ch.)

Null space $\Rightarrow$ $\dim$ Null(C)

$\dim$ Null$(C) = n - p$

Let Z be basis for null(C).

$\Rightarrow y \in$ Null$(C) \Leftrightarrow y = Zt$

Z: $n \times (n-p)$ the mat.

$\Rightarrow CLS \Leftrightarrow$

min $\|b - A(x_0 + Zt)\|_2^2$

s.t. Nothing !