Convexity ① $f(\lambda x_1 + (1-\lambda)x_2) = \lambda f(x_1) + (1-\lambda) f(x_2) \;\forall\; \lambda \in [0,1]$

② epigraph convex ③  contains chord joining 2 points

To Prove

① max of affines (pointwise)

② affine transformations

convex (affine) ⇒ convex

f(affine) ⇒ convex

$\min_x f_0(x)$

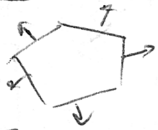
$f_i(x) \le 0 \quad i=1,\dots,m$

$h_j(x) = 0 \quad j=1,\dots,n$

$(x_j^2 = 1 \quad , \; j=1,\dots,p)$

min → max negate convex!

LP $\min c^T x$ s.t. $Ax \le b$, $Cx = d$



conic form

$\min c^T x : Ax = b$, $x \ge 0$, $x \in K$ (all variables ≥ 0)

QP

$\min c^T x + x^T Q x$

s.t. $Ax \le b$, $Cx = d$

$Q = Q^T$ is PSD/PD

QCQP

$\min c^T x + x^T Q x$

s.t. $c_i^T x + x^T Q_i x \le b_i$, $i=1,\dots,m$

$Cx = d$

SOCP

$\min c^T x$

s.t. $\|A_i x + b_i\|_2 \le c_i^T x + d_i$, $i=1,\dots,m$

$Cx = d$

- ℓ_∞ norm as constraints

↳ greater than ∀ components

- with booleans

↳ express as a ton of linear constraints

(otherwise Lagrange)

Minimax Inequality

$\min_{x \in X} \max_{y \in Y} L(x,y) \ge \max_{y \in Y} \min_{x \in X} L(x,y)$

Duality (weak)

$L(x,y) = \frac{1}{2}\|x\|_2^2 + y^T(Ax-b)$ $(y \ge 0)$

* scalars don't matter much

$p^* = \min_x \frac{1}{2}\|x\|_2^2, \; Ax \le b$

(y=dual variable)

$L(x,y) = f_0(x) + \sum_{i=1}^m y_i f_i(x) = f_0 + y^T f(x)$

$L(x,y) = \min_x \max_y$

$p^* \ge d^* = \max_{\lambda \ge 0} g(y)$

$g(y,0) = \min_x L(x,y)$

Rotated Second Order

$\|x\|_2^2 \le 2yz, \; y \ge 0, \; z \ge 0 \Longleftrightarrow$

$\left\| \begin{bmatrix} x \\ \frac{1}{\sqrt{2}}(y-z) \end{bmatrix} \right\|_2 \le \frac{1}{\sqrt{2}}(y+z)$

$w = (x, (y-z)/\sqrt{2})$

$t = (y+z)/\sqrt{2}$

$|u(t)| \le v$

$-v \le u(t) \le v$

* square root is not affine (unless considering const)

* bounding norms from below is not convex

$Ax \le b$

• is intersection of halfplanes

$S^T A S = \frac{1}{2}(S^T A S + S^T A^T S)$

booleans can also $[0,1]$

(Strong)

Slater's condition

① primal = convex

② dual convex

③ strictly feasible (not required for affine constraints)

Minimax Equality Theorem

(Sion's)

$p^* = \min_{x \in X} \max_{y \in Y} L(x,y)$

$p^* = d^* = \max_{y \in Y} \min_{x \in X} L(x,y)$

(i) $L(x,\cdot)$ concave

(ii) $L(\cdot,y)$ convex

(iii) X or Y are compact

e.g. $Y = \{y : \|y\| \le 1\}$

(iv) X & Y are convex

quadratic

$x^T Q x + c^T x \le t$

$w^T w \le 2yz$,

$z = \frac{1}{2}$ $w = Q^{1/2} x$,

$y = t - c^T x$

$\left\| \begin{bmatrix} \sqrt{2}\, Q^{1/2} x \\ t - c^T x - \frac{1}{2} \end{bmatrix} \right\|_2 \le t - c^T x + \frac{1}{2}$

or ×2 there

$\left\| \begin{bmatrix} w \\ \frac{y-z}{2} \end{bmatrix} \right\|_2 \le \frac{y+z}{2}$

or $\div \sqrt{2}$

Bilinear

• hold x, linear in y

• hold y, linear x

Linear can be concave or convex

* mechanism is square both sides to get back original constraint ⇒ squaring is affine

Primalization

if $\forall \lambda, \nu$ argmin $L(x,\lambda,\nu)$ is a singleton (unique) $\{x^*(\lambda,\nu)\}$ and if (λ^*, ν^*) are optimal for dual problem then $x^*(\lambda^*,\nu^*)$ is optimal for primal problem

Do 3d

Do 2

Linear Algebra

angle b/w vectors $\cos\theta = \frac{x^T y}{\|x\|_2 \|y\|_2}$

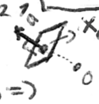
Epigraph $t \geq f(x)$

First order approximations $f(x_0) + \nabla f(x_0)^T (x - x_0) \approx f(x)$

$\nabla f(x) = \left[\frac{\partial f(x)}{\partial x_1} \; \dots \; \frac{\partial f(x)}{\partial x_n}\right]$

Line $\{x_0 + tu : t \in \mathbb{R}\}$

Hyperplane $H = \{x : a^T x = b\} \Rightarrow \{x : a^T(x - x_0) = 0\}$



$a^T x \geq b$ (acute w/ a)
$a^T x < b$ (opposite)

* center projection map
$f(x) = u^T(x - \hat{x})$ $\hat{x} = \text{ave}$

Chain Rule for gradient
$\left[\frac{\partial g_1(x)}{\partial x_j} \; \dots \; \frac{\partial g_m(x)}{\partial x_j}\right] \nabla f(g(x))$

Norms $\|x\|_2 = \sqrt{x^T x}$ $\{\text{card}(x)$

$\|x\|_1 = \sum_{i=1}^{n} |x_i|$ $\{\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$

$Ab = (Ab_1, Ab_2 \dots) = \begin{pmatrix} a_1^T B \\ \vdots \\ a_n^T B \end{pmatrix}$

$A = (A_1\ A_2)$, $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \Rightarrow Ax = A_1 x_1 + A_2 x_2$

$(A_1\ A_2)\begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = A_1 B_1 + A_2 B_2$ $\quad Tr = \sum_{i=1}^{n} A_{ii}$

$\begin{pmatrix} A_1 \\ A_2 \end{pmatrix}(B_1\ B_2) = \begin{pmatrix} A_1 B_1 & A_1 B_2 \\ A_2 B_1 & A_2 B_2 \end{pmatrix}$

$\langle A, B \rangle = Tr(A^T B) = \sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij} B_{ij} = \langle B, A \rangle$

$\nabla^2 f(x_0) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \dots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}_{x = x_0} = H$

Gradients

(i) $\alpha = y^T A x$
$\frac{d\alpha}{dx} = y^T A$
$\frac{d\alpha}{dy} = x^T A^T$

(ii) $\alpha = x^T A x$
$\frac{d\alpha}{dx} = x^T(A + A^T)$
if A symmetric
$= 2x^T A$

Cauchy-Schwartz

1) $x^T y \leq \|x\|_2 \|y\|_2$

2) $\max_{u : \|u\|_2 \leq 1} w^T u = \|w\|_2$ (where $u^* = \frac{w}{\|w\|_2}$)

Projections on a line

$\min_t \|x - x_0 - tu\|_2$

closed form

$t^* = u^T(x - x_0)$ (solve by squaring both sides)

When $\|u\|$ not norm

$z^* = x_0 + \frac{u^T(x - x_0)}{u^T u} u$

(else just numerator)

Projection on Hyperplane

$a^T z = b$

$\min_x \|x\|_2 : x \in H$

is b (b/c point closest to origin)

second order

$\dots + \frac{1}{2}(x - x_0)^T \nabla^2 f(x_0)(x - x_0)$

Holder
$\left(\frac{1}{p} + \frac{1}{q} = 1\right)$
$|x^T y| \leq \sum_{k=1}^{n} |y_k x_k| \leq \|x\|_p \|y\|_q$

Gram Schmidt

- pick one
- normalize
- project & subtract for $\perp$
- repeat

Fundamental Thm Lin Alg

$N(A) \perp R(A^T)$
$R(A) \perp N(A^T)$

Orthogonal

$U^T U = I$ $\quad \|Ux\|_2^2 = \|x\|_2^2$

$\cos\theta = x^T y$ $\quad \cos\theta' = (Ux)^T U y = x^T y$

- preserves lengths & angles

Inverses

$AA^{-1} = A^{-1}A = I_n$
$(AB)^{-1} = B^{-1}A^{-1}$
$(A^T)^{-1} = (A^{-1})^T$
$\det A = \det A^T$
$= \frac{1}{\det A^{-1}}$

Dyad uv^T

Matrix Norms

1) Frobenius $\|M\|_F = \sqrt{\sum_{i,j} M_{ij}^2}$
$= tr(M^T M)$

2) $\|M\|_2 = \max_{x \neq 0} \frac{\|Mx\|_2}{\|x\|_2} = \max_{\|x\|_2 = 1} \|Mx\|_2$ (largest singular value norm)

Full column rank & left inverse

$BA = I$
$B = (A^T A)^{-1} A^T$

Full row rank & right inverses

$AB = I$
$B = A^T(AA^T)^{-1}$

Rayleigh quotient

$\lambda_{min}(A) \leq \frac{x^T A x}{x^T x} \leq \lambda_{max}(A)$

Covariance matrix

$(w^T \hat{x} = \hat{s})$

$\Sigma = \frac{1}{m}\sum_{i=1}^{m}(x^{(i)} - \hat{x})(x^{(i)} - \hat{x})^T$

use decomposition to do analysis

$\sigma^2 = \sum_{i=1}^{m}(w^T x^{(i)} - \hat{s})^2 = \sum_{i=1}^{m}(w^T(x^{(i)} - \hat{x}))^2 = w^T \Sigma w$

Spectral

$\sum_{i=1}^{n} \lambda_i u_i u_i^T = U \Lambda U^T = A$

A symmetric

PSD $x^T A x \geq 0 \;\; \forall x \in \mathbb{R}^n$ $(\lambda \geq 0)$

$A = B^2$, $B \succeq 0$

(B nonsingular) $A = B^T B$

PD $x^T A x > 0 \;\; \forall 0 \neq x \in \mathbb{R}^n$ $(\lambda > 0)$

$A = B^2$, $B \succ 0$

PCA

- project on each singular vector one-by-one to get k rank approximation

SVD EVD=SVD is PD/PSD

$A = U \Sigma V^T$ (U & V col. = sing. values)

$M^T M v_j = \sigma_j^2 v_j$
$M M^T u_j = \sigma_j^2 u_j$
$\Rightarrow$ if M is sym $\sigma = |\lambda|$ & $V = \text{sgn}(\lambda_i)$

K-rank approx

1) SVD
2) Pick K largest sing values

U, V with σ = approx

Ellipsoids $\mathcal{E} = \{x \in \mathbb{R}^n : x^T P^{-1} x \leq 1\}$ $\quad P \succ 0 \Rightarrow P^{-1} \succ 0$

$x^T A^T A x = \|Ax\|_2^2$ $\quad \mathcal{E} = \{x \in \mathbb{R}^n : \|Ax\|_2 \leq 1\}$

Least Squares

a) $Ax = y$ A not invertible

$U\Sigma(V^T x) = y \Rightarrow \Sigma(V^T x) = U^T y$

b) $\min_x \|Ax - y\|_2 \Rightarrow \|\Sigma V^T - U^T y\|_2$

* introduce ℓ_2 regularizer, can still use SVD
* introduce ℓ_1 norm & can't use SVD

When A is full column rank

$x^* = (A^T A)^{-1} A^T y$

Line through 2 points

x_0 & $x_d \Rightarrow x_d - x_0$
centers origin @ x_0

(projection) problem

$\min_t \|tx_d - x\|_2^2 \Rightarrow$
$t^2 - 2t(x^T x_d) + x^T x$ (etc.)

- solve for t^*

Regularized LS $\lambda > 0$; A is rand deficient

$\min_x \|Ax - y\|_2^2 + \lambda\|x\|_2^2$
$\rightarrow \min \|\tilde{A}x - \tilde{y}\|_2^2$

Power iteration

$\rightarrow$ HW8 123 $\Rightarrow$ approx/interpolate but make sure you maintain the values you know

MT $\|X - \hat{X}\|_F \Rightarrow \hat{x} = \begin{cases} M_{ij} \\ \hat{x}_{ij} \end{cases}$

Weighted Regularized LS

$\min \|Ax - y\|_2^2 + x^T W x$; $\quad x^* = (A^T A + W)^{-1} A^T y$

$\tilde{A} = \begin{pmatrix} A \\ \sqrt{\lambda} I_n \end{pmatrix}$ $\quad \tilde{y} = \begin{pmatrix} y \\ 0 \end{pmatrix}$

Minimum norm

$\min \|x\|_2 : Ax = y$
undetermined; inf sol

$S_{\bar{x}} = \{x : \bar{x} + z;\ z \in N_u(A)\}$

$x^* \perp$ to $N(A) \Rightarrow \in R(A^T)$

$Ax^* = y = AA^T \xi$
$x^* = A^T \xi$
$\xi = (AA^T)^{-1} y$

Linearly Constrained

$\min \|Ax - y\|_2^2$; $Cx = d$

$x = \{x_0 + Nz;\ z \in \mathbb{R}^k\}$ $\quad \tilde{A} = AN$

$\min_z \|\tilde{A}z - \tilde{y}\|_2$ $\quad \tilde{y} = y - Ax_0$

Discretize, first order

$\frac{\partial f}{\partial x}(x, y) \approx \frac{1}{h}(f(x + h, y) - f(x, y))$

1) Pick an increment that makes sense

$\nabla f(X_{ij}) \approx G_{ij} := \begin{pmatrix} k(\hat{F}_{i+1,j} - \hat{F}_{ij}) \\ k(\hat{F}_{i,j+1} - \hat{F}_{ij}) \end{pmatrix}$

Cholesky $A \succ 0$ (PD)

$L = [\Delta]$ but $A \succeq 0$ too

$A = LL^T$

Convexity

- constrain to 1D
$\rightarrow$ check variable possibilities
ex $|x - LR^T| \rightarrow |1 - LR| \rightarrow |1 - L^2|$ (not convex)

$\min\ kv + \sum_{i=1}^{n} \max(0, |x|_i - v) : Ax \leq b$
$\rightarrow + \sum s_i$ $\quad s_i \geq x_i - v$, $\quad s_i \geq -x_i - v$, $s_i \geq 0$

$\rightarrow$ write as convex obj.