Midterm study guide
10/09/2017

- BASICS

- notation
  - $x \preceq y$ ~ x less than y elementwise
  - $X \preceq Y$ ~ $Y-X$ is PSD ~ $\forall_v \; v^T X v \leq v^T Y v$
- Linearity
  - inner product: $tr(X^T Y)$
  - Cauchy-schwartz inequality:
    $\|f(ax+b)\| \leq \|f(a)\| \, \|f(b)\|$
    or
    $|x^T y| \leq \|x\|_2 \|y\|_2$

- MATRIX PROPERTIES

- nonsingular = invertible
- $x^T A x = tr(x x^T A)$
- inverse
  - orthogonal : $A^{-1} = A^T$
  - columns are orthonormal
  - $\|Ax\|_2 = \|x\|_2$
  - pseudo-inverse = Moore-Penrose inverse
    $A^+ = (A^T A)^{-1} A^T$

- MATRIX CALC

$\nabla^2 f(x)_{ij} = \dfrac{\partial^2 f(x)}{\partial x_i \partial x_j}$

- examples
  - $\nabla_x a^T x = a$
  - $\nabla_x x^T A x = 2Ax$
  - $\nabla_x^2 x^T A x = 2A$
  - $\nabla_x \log \det X = X^{-1}$

- NORMS

- defn $\Rightarrow$ convex
  1. nonnegative
  2. definite $f(x)=0$ iff $x=0$
  3. proportionality
  4. triangle inequality
- vector norms
  - L_p-norms: $\|x\|_p = \sqrt[p]{\sum |x_i|^p}$
    - L_0
    - L_1
    - L_∞ = Chebyshev norm
  - quadratic norms
    - P-quadratic norm: $\|x\|_P = (x^T P x)^{1/2} = \|P^{1/2} x\|_2$, $P \in S^n_{++}$
  - dual norm of $\|\cdot\|$
    $\|z\|_* = \sup\{z^T x \mid \|x\| \leq 1\}$
- matrix norms
  - frobenius norm - like l_2
  - sum-abs-value, max-abs-value
- operator norm
  - vector norms $\|\cdot\|_a$, $\|\cdot\|_b$
    $\|X\|_{a,b} = \sup\{\|Xu\|_a, \|u\|_b \leq 1\}$
    if a, b both Euclidean norms,
    spectral norm = L_2 norm = max singular value
    $\|X\|_2 = \sigma_{max}(X)$

Appendix

EIGENSTUFF

- when $A \in S$
  - $\det(A) = \prod_i \lambda_i$
  - $tr(A) = \sum_i \lambda_i$
  - $\|A\|_2 = \max |\lambda_i|$
  - $\|A\|_F = \sqrt{\sum_i \lambda_i^2}$
  - $\lambda_{max}(A) = \sup_{x \neq 0} \dfrac{x^T A x}{x^T x}$
  - $\lambda_{min}(A) = \inf_{x \neq 0} \dfrac{x^T A x}{x^T x}$
- PSD
  - usually symmetric
    $x^T A x \geq 0 \;\; \forall x$
    $\exists X$ s.t. $A = X^T X$ ~ assumes symmetric

DIAGONALIZATION

= eigenvalue decomposition
= spectral decomposition
if A symmetric,
$A = Q \Lambda Q^T$
~ Q colums eigenvecs, orthonormal
if X, Y symmetric,
$tr(YX) = tr(Y \sum \lambda_i q_i q_i^T)$
generalized eigenvalue decomp: for 2 symmetric matrices

SINGULAR VALUE DECOMP

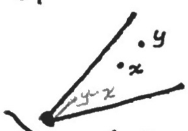
$A = U \Sigma V^T$
$\sigma \in \Sigma$ singular value = square roots of nonnegative eigenvalues of AA^T
when PD,
$\Sigma = \Lambda$, $U \Sigma V^T = Q \Lambda Q^T$
$A^+ = V \Sigma^+ U^T h$

MISC

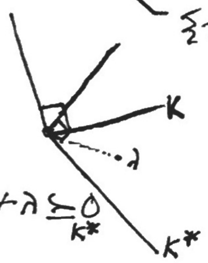
- Hadamard's determinantal inequality:
  $\det X = \min_{\text{o.n. basis } \{v_1, \ldots, v_n\}} \prod v_i^T X v_i$
- Schur complement of $X = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$
  $S = C - B^T A^{-1} B$
  $X \succeq 0 \Leftrightarrow S \succeq 0$

Ch 2 - Convex Sets

- affine set: $x_1, x_2 \in C, \theta \in \mathbb{R} \Rightarrow \theta x_1 + (1-\theta)x_2 \in C$
- affine hull: $\text{aff } C = \{ \sum \theta_i x_i \mid x_i \in C, \sum \theta_i = 1 \}$
- convex set: $0 \le \theta \le 1 \Rightarrow \theta x_1 + (1-\theta) x_2 \in C$
- cone: $\theta \ge 0 \Rightarrow \theta x \in C$
- operations that preserve convexity
  - intersection
  - pointwise max of affine funcs
  - composition
  - affine
  - perspective
  - linear fractional = projective
- generalized inequalities:
- proper cone K ~ convex, closed, pointed
  $x \preceq_K y \Leftrightarrow y - x \in K$



- separating hyperplane thm — C, D convex $C \cap D = \emptyset \Rightarrow \exists a \neq 0, b$ s.t. $a^T x \le b \ \forall x \in C$, $a^T x \ge b \ \forall x \in D$
- supporting hyperplane thm — $\{x \mid a^T x = a^T x_0\}$ where x_0 on boundary of convex C
- dual cone K^*
  $= \{ y \mid x^T y \ge 0 \ \forall x \in K \}$
  $\preceq_{K^*}$ is dual of $\preceq_K$
  $x \preceq_K y \Leftrightarrow \lambda^T x \le \lambda^T y \ \forall \lambda \succeq_{K^*} 0$



GEOMETRY

- ellipsoid: $\{ x \in \mathbb{R}^n \mid (x - x_c)^T P^{-1} (x - x_c) \le 1 \}$
  $\{ x_c + Au \mid \|u\|_2 \le 1 \}$
  ~ P symmetric + PSD
- hyperplane: $\{ x \mid a^T x = b \} \rightarrow$ creates halfspace
- norm cone: $\{ (x,t) \mid \|x\| \le t \}$
- polyhedron: $\{ x \mid Ax = b, Cx = d \}$
  $= \{ \sum_{i=1}^{k} \theta_i v_i \mid \sum_{i=1}^{m} \theta_i = 1, \theta_i \ge 0 \}$ $m \le k$
- simplex: $\text{conv}\{ v_{0:k} \}$

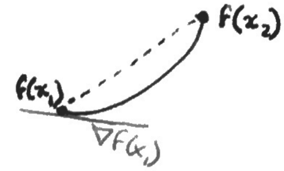
definitions

Ch 3 - Convex funcs

1. Jensen's inequality: $0 \le \theta \le 1$

$f(\theta x_1 + (1-\theta)x_2) \le \theta f(x_1) + (1-\theta) f(x_2)$

$f(E[X]) \le E[f(x)]$

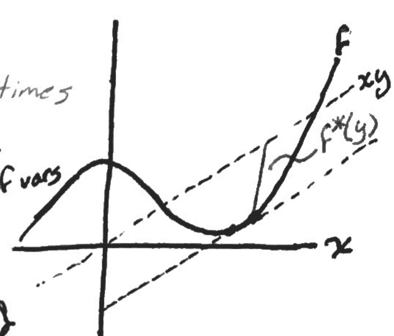


2. $\nabla^2 f(x) \succeq 0$

3. $f(x_2) \ge f(x_1) + \nabla f(x_1)^T (x_2 - x_1)$
   - can show this by restricting to an arbitrary line
4. consider epi f

- other concepts
  - epigraph $\text{epi } f = \{(x,t) \mid x \in \text{dom} f, f(x) \le t\}$
  - extended value extension
    $\tilde{f}(x) = \begin{cases} f(x) & x \in \text{dom } f \\ \infty & x \notin \text{dom } f \end{cases}$
  - wide sense function ~ can be $\pm\infty$
    $\text{dom } f = \{x \mid f(x) < \infty\}$
  - wide sense convex func:
    given convex set $F \subseteq \mathbb{R}^{n+1}$
    $f(x) = \inf\{t \in \mathbb{R} \mid (x,t) \in F\}$
  - α-sublevel set of convex func is convex

- operations that preserve convexity
  - nonnegative weighted sums ~ multiples for logs
  - affine map
  - pointwise max of convex
  - composition
  - perspective
  - minimization ~ sometimes

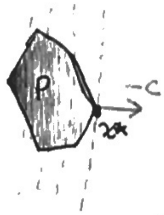


- conjugate of f ~ can use change of vars
  $f^*(y) = \sup_{x \in \text{dom} f} y^T x - f(x)$
  $\text{dom } f^* = \{y \mid f^*(y) \text{ is finite}\}$
  - called Legendre transform when f differentiable
- Fenchel's inequality: $f(x) + f^*(y) \ge x^T y$
- $f^{**} = f$ iff convex, closed
- ex. $f(S) = \log\det X^{-1}$
  $f^*(Y) = \sup_X [\text{tr}(YX) + \log\det X]$
  $= -n - \log\det(-Y)$ if $-Y \in S^n_{++}$
- ex. $(u \log u)^* = e^{v-1}$
- ex. $\|\cdot\|^* = \begin{cases} 0 & \|y\|_* \le 1 \\ \infty & \text{oth.} \end{cases}$
- can use conj. to go other way: $f(y) = \sup_x (y^T x - f^*(x))$

...MIZATION

...ndard form:

$p^* =$ minimize $f_0(x)$
s.t. $f_i(x) \le 0$
$h_i(x) = 0$



- equivalent problems
  - change of vars
  - constraint transformations
  - slack vars
  - eliminating equalities
  - eliminating linear equalities
  - introducing equalities
  - optomizing over some vars ~ ex. quadratic
  - epigraph form: min t s.to $f_0 \le t$
  - implicit + explicit constraints

CONVEX OPTIMIZATION

minimize $f_0(x)$
s.t. $f_i(x) \le 0$ $i = 1:m$
$a_i^T x = b_i$ $i = 1:p$ ~ in class only assumed convex

where $f_{0:m}$ convex

- x optimal if
  1. x feasible
  2. $\nabla f_0(x)^T (y - x) \ge 0$ $\forall y$ feasible
- if unconstrained, $\nabla f_0(x) = 0$
- if equality only $Ax = b$: $\nabla f_0(x) \perp \mathcal{N}(A)$
- $x \succeq 0$: $\nabla f_0(x) \succeq 0$; $x_i (\nabla f_0(x))_i = 0$

optimality criteria

- equivalent convex problems
  - eliminating equality constraints
  - introducing equality constraints
  - slack vars ~ for linear inequalities
  - epigraph form
  - minimizing over some vars

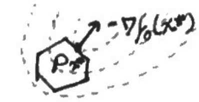
4 - Convex Optimization Problem

LINEAR OPTIMIZATION

minimize $c^T x + d$
s.t. $Gx \preceq h$
$Ax = b$

- standard form: $x \succeq 0$ is the only inequality
- standard dual: max $-b^T \nu$ s.t. $A^T \nu + c \ge 0$
- Linear-fractional program

minimize $\dfrac{c^T x + d}{e^T x + f}$ ~ can be converted to LP
s.t. $Gx \preceq h$
$Ax = b$



QUADRATIC OPTIMIZATION

minimize $\frac{1}{2} x^T P x + q^T x + r$
s.t. $Gx \preceq h$
$Ax = b$
where $P \in S_+^n$

- QCQP ~ inequality constraints also convex
- ex. minimize $\|Ax - b\|_2^2$

5 - Duality

der minimize $f_0(x)$

s.t. $f_i(x) \le 0$

$h_i(x) = 0$

- Lagrangian: $L(x, \lambda, \nu) = f_0(x) + \sum \lambda_i f_i(x) + \sum \nu_i h_i(x)$
- Dual function: $g(\lambda, \nu) = \inf_{x \in D} L(x, \lambda, \nu)$ ~ g always concave
  - $\lambda \succeq 0 \Rightarrow g(\lambda, \nu) \le p^*$
- (λ, ν) dual feasible if
  1. $\lambda \succeq 0$
  2. $(\lambda, \nu) \in \text{dom } g$
  - when $p^* = -\infty$, dual infeasible
  - when $d^* = \infty$, primal infeasible
- dual related to conjugate func
  - ex. min $f(x)$ s.t. $x = 0 \Rightarrow g(\nu) = -f^*(-\nu)$
- Lagrange dual problem: maximize$_{\lambda, \nu}$ $g(\lambda, \nu)$ s.t. $\lambda \succeq 0$
- weak duality: $d^* \le p^*$
  - optimal duality gap: $p^* - d^*$
- strong duality: $d^* = p^*$ ~ requires more than convexity
- Slater's condition ~ if problem convex + met $\Rightarrow$ strong duality

  $\exists x \in \text{relint } D$

  $f_i(x) < 0$, $i = 1:m$, $Ax = b$ ~ point is strictly feasible

  ~ to weaken this, affine f_i can be ≤ 0
- Sion's minimax thm: $x \to f(x, y)$ ~ conditions

  $\Rightarrow \min_x \sup_y f(x, y) = \sup_y \min_x f(x, y)$

OPTIMALITY CONDITIONS

- duality gap $f_0(x) - g(\lambda, \nu)$
  - can use stopping condition duality gap $\le \varepsilon_{abs}$ to be ε_{abs}-suboptimal
- Strong duality yields complementary slackness

  $\lambda_i^* f_i(x^*) = 0 \quad i = 1:m$
- KKT optimality conditions ~ assume f_0, f_i, h_i differentiable ~ assume strong duality
  1. $f_i(x^*) \le 0 \quad i = 1:m$
  2. $h_i(x^*) = 0$
  3. $\lambda_i^* \ge 0$
  4. $\lambda_i^* f_i(x_i^*) = 0$
  5. $\nabla f_0(x^*) + \sum \lambda_i^* \nabla f_i(x_i^*) + \sum \nu_i^* \nabla h_i(x^*) = 0$
- if primal problem convex, ~~strong duality~~ holds + sols to KKT are primal + dual optimal with zero duality gap

THMs of ALTERNATIVES

- weak alternative - at most one of 2 is true
- strong alternative - exactly one is true
  - ex. Fredholm alternative
  - ex. Farkas's Lemma
    1. $\exists x \quad Ax \le 0 \wedge c^T x < 0$
    2. $\exists y \quad y \succeq 0, \ A^T y + c = 0$