Final Exam Review

Sunday, May 4, 2025 11:01 PM

Key Concepts

Convex Sets (chapter 2)

- Basic definition → x_1 and $x_2 \in C \Rightarrow \theta x_1 + (1-\theta)x_2 \in C$; $\theta \in [0,1]$
- Relationship to
    - Lines/line segments
    - Affine sets ← $\sum_{i=1}^{n} \theta_i = 1$
    - Subspaces → $x_1, x_2, \ldots, x_n \Rightarrow \theta_1 x_1 + \cdots + \theta_n x_n$
- Cones and Convex Cones
- Hyperplanes → $f(x) = a^T x + b \Rightarrow \dim(f(x)) = n-1$
- Half spaces → $a^T x + b \leq 0$ or $a^T x + b \geq 0$
- Ellipsoids → $x^T Q x = a$; $Q \succ 0$
    - $Q = \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix}$; $1/\sqrt{8}$, $\frac{1}{\sqrt{2}}$
- Norm cone and second-order cone
- Polyhedra / polytopes
- Positive semidefinite cone
- Hulls
    - Affine hull
    - Convex hull
    - Conic hull
- Operations that preserve convexity
    - (infinite) intersection
    - Affine functions
        - Image of a convex set
        - Preimage (inverse image) of a convex set
    - Image of a convex set under the perspective function

Convex functions (chapter 3)

- Basic (Jensen's inequality-based) definition → $f(\theta x_1 + (1-\theta)x_2) \leq \theta f(x_1) + (1-\theta) f(x_2)$ $\forall\ \theta \in [0,1]$
- First-order condition
- Second-order condition
- "Restriction to an arbitrary line" condition ☆ → $\log\det(X)$, $(\det(X))^{1/n}$, $X \succ 0$
    - computation of a gradient
    - Computation of a Hessian
    - Computation of a directional derivative
- Epigraph definition
- Strict convexity
- Lipschitz continuous gradients (smooth functions)
- Strong convexity
    - Global minima
    - Local minima
    - Saddle points
    - → Stationary points
    - Mapping to concave functions

- Local minima
- Saddle points
- Mapping to Concave functions

- Geometrical understanding
  - Uniform linear lower bound
  - Uniform quadratic upper bound
  - uniform quadratic lower bound
  - Quadratic upper bound on a compact set
- Sublevel set of a convex function
- Some common convex / concave functions
- operations that preserve convexity
  - Nonnegative weighted sums
  - Composition with an affine map
  - Point wise maximum and Supremum
  - Composition of functions
  - infimum over a convex set

Convex optimization (chapter 4)

- Standard form
- Terminology
  - Domain of the problem
  - feasible set
  - Infeasible problem
  - Optimal point and optimal value
  - local optimality
  - Equivalent problems
  - slack variables

$\boxed{A}x = 0$

$Ax \preceq 0$

- Eliminating linear equality constraints
- Necessary and sufficient condition for global optimality of a convex problem
  - Unconstrained problem $\Leftrightarrow \nabla f(x^*) = 0$
  - Constrained problem
- Common convex optimization problems
  - Linear program
  - Quadratic program
  - Quadratically constrained quadratic program
  - Second-order cone program
  - Semidefinite program

Duality (chapter 5) $p^* \Rightarrow$ primal optimal value

- Lagrangian of a constrained optimization problem
- Dual function
  - Conjugate of a function $\longrightarrow$ $d^* \Rightarrow$ dual optimal value
- Dual problem
  - Dual feasibility

$d^* \leq p^*$

- Dual problem — $d^* \Rightarrow$ dual optimal value
  - Dual feasibility
  - weak duality — $d^* \leq p^*$, $g(\lambda, \nu) \leq p^*$, $\lambda \succeq 0$
  - Strong duality — $d^* = p^*$
  - Duality gap — $p^* - d^*$, $f_0(x) - g(\lambda, \nu)$
- Slater's condition for strong duality
- KKT conditions for strong duality
  - Necessary conditions
  - Sufficient conditions
  - Necessary and sufficient conditions

$\arg\min_x L(x, \lambda^*, \nu^*) = x^*$

Unconstrained optimization algorithms (chapter 9)

- Condition number of a problem
- Gap to optimality as a function of gradient norm — strongly convex
- Descent methods
- Gradient descent
  - Range of step sizes
  - Optimal fixed step size — $1/L$
  - Backtracking line search for step size
  - Convergence rates
  - Dependence on the condition number

$f(x) - p^* \leq \frac{1}{2m} \|\nabla f(x)\|_2^2$ — strong convexity

- Steepest descent
  - Dual norm
  - P-Quadratic norm — $\Delta x = -P^{-1} \nabla f(x)$
  - Interpretation as preconditioning
  - Convergence rates
- Newton's method
  - Full Newton step — $\Delta x = -(\nabla^2 f(x))^{-1} \nabla f(x)$
  - Relation to steepest descent
  - Newton decrement — stopping criterion
  - Invariance to affine transformation
  - Independent w.r.t. K
  - Two phases of convergence
    - Linear
    - Quadratic

Constrained Optimization with Equality Constraints (chapter 10)

- Solution of a Quadratic program
- General convex problems
  - Eliminate equality constraints
  - Solve the dual problem
  - Approximate as a quadratic program
- Feasible start Newton's method

— Approximate as a quadratic program

— Feasible start Newton's method

— Infeasible start Newton's method

— Backtracking line search w.r.t. the primal and dual residual vectors

$$\begin{bmatrix} P & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} z \end{bmatrix} = \begin{bmatrix} c \end{bmatrix}$$

P is Symmetric

unknown

$$\begin{bmatrix} -q \\ b \end{bmatrix}$$