1 9/6: Linear least squares - review

- Least squares problems
- Derivation of the normal equations
- Moore-Penrose pseudoinverse
- When is a least squares system solvable?
- What is the gradient? What is the Jacobian? How do they relate?
- Taylor expansions
- Different ways of computing the gradient:
  - Directly compute partial derivatives
  - Use a Taylor expansion
  - Use known facts from matrix calculus

2 9/8: Review of multivariable calculus

- Definition of ordinary derivative
- Partial derivatives
- The gradient
- The gradient points in the direction of steepest ascent
- The Jacobian
- Difference between row vectors and column vectors
- The Hessian
- Multivariable Taylor expansions
- Big O notation (as used in analysis)

3 9/13: Nonlinear least squares

- Nonlinear least squares problem
- How to compute the Jacobian of the nonlinear least squares cost function
- Quasi-Newton approximation used for Gauss-Newton

4 9/15: Newton's method intro - Gauss-Newton

- Newton's method in 1D - rootfinding
- What is an iterative scheme?
- Derivation of Newton's method in 1D for rootfinding
- Two versions of multidimensional Newton's method:
  - For "Rootfinding" (solving nonlinear systems)
  - For minimization
- Derive the two different versions using Taylor expansions
- Derive the Gauss-Newton iteration

5 9/20: Gradient descent

- Computational complexity of Newton's method for minimization
- The Newton step and the Newton system
- Recall: LU decompositions and their costs
- When is the cost of Newton's method justified?
- Descent direction - motivation, definition, intuition
- Level sets of cost functions and relation to cost function gradient
- The negative gradient is always a descent direction
- Motivation for why we need line search - simple example
- Exact line search

6 9/22: Behavior of gradient descent

- Behavior of gradient descent with exact line search applied to quadratic form
- Newton's method finds stationary point of a quadratic form in one step - simple example
- Gradient of a quadratic form - applying Newton's method

7 9/27: Optimality - convexity

- Review of first-order necessary conditions in 1D
- Local minima - strict local minima - (counter)examples
- Global minima - strict global minima
- Necessary and sufficient conditions
- Definiteness of matrices - spectral theorem
- Convexity - strict convexity - relationship to optimality
- Relationship between convexity and definiteness of Hessian

8 9/29: Newton's method revisited

- Another derivation of Newton's method
- Newton's method may not converge
- Rate of convergence of a sequence
- Theorem: quadratic convergence of Newton's method
- Motivation for quasi-Newton methods
- Modified Newton's method

9 10/4: Line search methods

- The Armijo condition/sufficient decrease condition - intuition
- The angle condition
- Gradient-related directions
- Lipschitz continuity
- Theorem: convergence of gradient-related iteration with line search to local minimum

10 10/6: Quasi-Newton methods

- Secant method in 1D
- The secant condition
- Interpretation of the secant condition
- Derivation of SR1 quasi-Newton method
- General algorithm for a quasi-Newton method
- No rank 1 quasi-Newton method which is both symmetric and positive definite
- The Broyden class: special cases = BFGS and DFP

11 10/13, 10/18: More Quasi-Newton

- Recall: the singular value decomposition
- Recall: the Cholesky decomposition
- Rank-1 updates for a Cholesky factorization
- Givens rotations
- Sherman-Morrison formula
- Accumulating the inverse Hessian in a quasi-Newton method

12 10/25: Geometry of LPs

- General form of a constrained optimization problem
- What is an LP?
- Simple example of an LP in 2D
- Basic properties and geometry in 2D and 3D
- A simple algorithm for solving an LP in few dimensions
- Matrix form of linear inequality constraints
- Computational complexity of direct solution of an LP
- Example: the transportation problem
- Reducing LPs to standard form

13 10/27: Standard form of an LP

- LP in standard form
- Any LP can be converted to standard form
- Recipe for converting an LP to standard form
- Examples - intuition

14 11/3: Extreme points and basic solutions

- Recall: rank-nullity theorem
- Affine dimension of the feasible set of an LP
- Extreme points - basic solutions - intuition - examples
- Theorem: for LPs, extreme points = basic solutions

15 11/8: Intro to KKT conditions

- Recall: method of Lagrange multipliers from Calc III
- Gradient of active constraint function is normal to boundary
- Geometric motivation of Lagrange multipliers
- Combine unconstrained case and constrained case using KKT conditions
- Active set - constrained local minima
- Theorem: KKT conditions are constrained first-order necessary conditions
- Derive KKT conditions for LPs

16 11/10: More KKT

- Linear independence constraint qualification (LICQ)
- KKT conditions for standard form LP
- KKT conditions are *sufficient* for LPs - proof

17 11/15: The dual LP

- Heuristic motivation for dual LP based on consideration of KKT conditions for primal LP
- KKT conditions for dual LP
- KKT conditions of dual LP = KKT conditions of primal LP
- Motivation of the concept of a dual problem - primal and dual LPs are dual
- Weak duality - strong duality
- Theorem: LPs satisfy strong duality

18 11/17: Interior point methods for LPs

- Duality gap for LP
- Relaxed KKT conditions for LP
- Applying Newton's method to KKT conditions
- Need to keep iterations strictly feasible
- Algorithm: primal-dual interior point method for solving LP

19 11/22: Lagrangian duality

- Relaxing constraints using indicator functions
- The dual function - the dual problem
- Weak duality (max of dual function $\leq$ min of primal function)
- Saddle point condition
- Theorem: strong duality $\iff$ saddle-point condition holds

20 11/29: Optimality conditions for linear equality constraints

- Feasible direction - constrained local minima

- Nonlinear minimization problem with linear equality constraints - characterization of feasible directions for this type of problem
- Recall: QR decomposition
- Computing the basis for the nullspace of the constraint matrix
- The reduced gradient - the reduced Hessian
- Necessary conditions - sufficient conditions
- Connection with Lagrange multipliers - interpretation of Lagrange multipliers for linear equality constraints
- Sensitivity - another interpretation of Lagrange multipliers

21 12/6: Optimality conditions for nonlinear constraints

- Regular points for nonlinear equality constraints
- Regular points for nonlinear inequality constraints
- Characterization of regular points
- Example: two equality constraints
- Example: single inequality constraint
- Constraints defining a feasible set aren't unique
- Necessary and sufficient conditions for nonlinear equality constraints
- Nondegenerate constraints
- Necessary and sufficient conditions for nonlinear inequality constraints

22 12/8: Algorithms for nonlinear constrained problems

- Recall: basic approach to solving unconstrained problems
- Compact form of KKT conditions for nonlinear equality constraints
- Problems with solving nonlinear system for nonlinear equality constraints using Newton's method:
  - Convergence only in basin of attraction

- Iteration can diverge
- No guarantee that iterates are feasible
- Can't use line search to drive convergence to a local minimum - can't use L to track progress

- Indefiniteness of Lagrange function
- $L = f$ for feasible points
- An idea for a more general projection method
- Other ways of solving: interior point methods - merit functions
- Example: ℓ_2 regularization