Review Guide for In-Class Midterm Exam on *Monday, 24 October 2016*

The Exam is closed book. *However, you may bring your own notes on a single sheet of letter-sized (*8.5 *inch by* 11 *inch) or smaller paper. Such notes must be entirely your work!*

The first Midterm Exam will cover Chapters 1, 2, 3, and parts of Appendices A.1 and A.2 in Nocedal & Wright, *Numerical Optimization*, 2nd edition. The first material below states what is excluded material that will *not* be on the Exam. After that I state the *specific* material that will be covered. My goal is to only cover topics that have appeared on homework and in lecture.

The problems will be of these types: state definitions, prove or show theorems/lemmas which follow reasonably-directly from the definitions or from known facts, give examples with certain properties, or describe or illustrate/sketch certain concepts and examples.

Because you will not be able to refer to the book during the exam, so I will not ask you to "state theorem 2.1" or anything like that. The only exception is that I will refer to problems in "form (1.1)" as short-hand for the standard (general) continuous optimization problem.

Please get together with other students and work through this Review Guide. Be honest with yourself about what you do and don't know, and talk it through and learn!

Excluded material. The following material will *not* be on the exam:

- the "transportation problem" on pages 5–6
- any "trust region" methods, such as discussions on pages 19–20 and 25–26
- the "symmetric-rank-one (SR1)" formula (page 24)
- *strong* Wolfe conditions (equations (3.7) page 34)
- Goldstein conditions (page 36)
- "line search algorithm for the Wolfe conditions" (pages 60–62)
- section 3.4 on "Newton's method with Hessian modification" (pages 48–49)
- those parts of section 3.5 on "a line search for the Wolfe conditions" (pages 60–62)
- in Appendix A.1:
  - the dual norm, equation (A.6) (page 601)
  - the SVD (page 604)
  - material on determinant and trace (pages 605–606)
  - the QR and the symmetric indefinite factorizations (page 609–612)
  - the interlacing eigenvalue theorem, error analysis and floating-point arithmetic, and conditioning and stability (pages 613–617)
- in Appendix A.2:
  - the topology of Euclidean space $\mathbb{R}^n$ and convex sets in $\mathbb{R}^n$ (pages 620–623)
  - the implicit function theorem (pages 630–631)

Definitions and Notation. Be able to state and use the definition, and/or use the notation/language correctly:

- problem (1.1) on page 3:
  - objective function and constraint functions
  - equality and inequality constraints
  - constrained versus unconstrained problems
- convex *set* (page 8)
- convex and strictly-convex *function* (page 8)
- global and/or local minimizer (page 12)
- strict and/or isolated local minimizer (page 13)
- gradient $\nabla f(x)$ and Hessian $\nabla^2 f(x)$ (page 626)
- directional derivative $\nabla f(x)^\top p$ (pages 628–629)
- vector norms (A.2)–(A.5) (page 600)
- matrix norms (A.7)–(A.10) (page 601)
- condition number (A.11) (page 601)
- symmetric, positive-definite, and positive-semidefinite matrices (page 599)
- short-hand: SPD = symmetric and positive-definite
- triangular, diagonal, and nonsingular matrices (page 599)
- eigenvalues and eigenvectors (page 603)
- Euclidean matrix norms $\|A\|$, $\|A^{-1}\|$ of SPD matrices from eigenvalues (page 605)
- descent direction: $\nabla f(x_k)^\top p_k < 0$ (page 21)
- line search: find α_k by solving or approximately solving (2.10) (page 19)
- Wolfe conditions (3.6a), (3.6b) (page 34)
- Q-linear, Q-superlinear, and Q-quadratic convergence in $\mathbb{R}^n$ (pages 619–620)

Algorithms. Be able to state the algorithm. Be able to illustrate with an example or give a sketch.

- steepest descent method: $p_k = -\nabla f(x_k)$ & $x_{k+1} = x_k + \alpha_k p_k$ (page 21)
- Newton method: $p_k^N = -\nabla^2 f(x_k)^{-1} \nabla f(x_k)$ & $x_{k+1} = x_k + \alpha_k p_k^N$ (page 22)
- quasi-Newton method: $p_k = -B_k^{-1} \nabla f(x_k)$ & $x_{k+1} = x_k + \alpha_k p_k$ (page 24)
- BFGS update formula (2.19) gives a quasi-Newton method (page 24)
- backtracking line search (page 37)
- triangular forward/back substitution for $Ly = b$ or $Ux = y$ (pages 606–607)
- Gaussian elimination as LU factorization/decomposition $A = LU$ or $PA = LU$ (pages 606–607)
- Cholesky factorization $A = LL^\top$ for SPD matrix A (pages 608–609)

Theorems and Lemmas. Understand these and remember them as facts. Be able to illustrate with an example or give a sketch. You may use these facts as needed in proving other propositions, but mention such use.

- Taylor's Theorem, especially (2.4) and (2.6) (page 14)
- *Theorem* (spectral decomposition): If $A \in \mathbb{R}^{n\times n}$ is symmetric then there exist real eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and an orthonormal basis of eigenvectors $q_1, q_2, \ldots, q_n$ so that $Aq_i = \lambda_i q_i$ and $A = Q\Lambda Q^\top$ where $Q = \big[q_1\big|\ldots\big|q_n\big]$ and Λ is diagonal with entries λ_i. (page 604)
- first-order necessary conditions Theorem 2.2 (page 14) [***be able to prove***]
- second-order necessary conditions Theorem 2.3 (page 15)
- second-order sufficient conditions Theorem 2.4 (page 16)
- Theorem 2.5 on minimizers of convex functions (page 16) [***be able to prove***]
- Sherman-Morrison-Woodbury formula (A.27) (page 612)
- Lemma 3.1: there exist step sizes α_k so that Wolfe conditions are satisfied (page 35)
- Theorem 3.2: if p_k are descent directions, f is bounded below, ∇f is Lipschitz, and the step sizes α_k satisfy the Wolfe conditions then $\sum_{k=0}^{\infty}(\cos\theta_k)^2\|\nabla f(x_k)\|^2 < \infty$ (pages 38–39)
- Corollary: steepest descent with a Wolfe-satisfying line search gives x_k so that $\|\nabla f(x_k)\| \to 0$ (pages 39–40) [***be able to prove***]
- Theorem 3.3: steepest descent on a quadratic function $f(x) = \frac{1}{2}x^\top Qx - b^\top x$ gives at least linear convergence in Q-norm (page 43)
- Theorem 3.5: Newton's method gives quadratic convergence under reasonable conditions, once you are close enough to the minimizer x^* (page 44)