CNAc: Continuous Optimization
Revision questions

Honour School of Mathematics, Oxford University
Trinity Term 2006, Dr Nick Gould

Here is a long list of questions which summarize the material taught in the course. Once you have finished revising, see how many of these you know the answers to *without* looking at your notes.

Part 1: optimality conditions

1.1. What does it mean for a function $\mathcal{F} : \mathcal{X} \rightarrow \mathcal{Y}$ to be Lipschitz continuous throughout the open set $\mathcal{X}$?

1.2. Give a bound on the error when the Lipschitz-continuously-differentiable function f is approximated by its first-order (linear) Taylor's approximation around x at the point $x + s$. What would the error be if a second-order (quadratic) Taylor approximation were used—what extra condition is required?

1.3. State sufficient optimality conditions for x_* to be an isolated local minimizer of the twice-continuously differentiable function $f(x)$.

1.4. Give first- and second-order necessary optimality conditions for x_* to minimize $f(x)$ when $c(x) = 0$. What assumptions do you need? Explain briefly what a constraint qualification is, and give two general examples where it is automatically satisfied.

1.5. State Farkas lemma, and explain it in words.

1.6. Give first- and second-order necessary optimality conditions for x_* to minimize $f(x)$ when $c(x) \geq 0$. What assumptions are you making? Do the same for the problem of *maximizing* $f(x)$ when $c(x) \leq 0$.

Part 2: Line-search methods for unconstrained minimization

2.1. What doe we mean by a descent direction for $f(x)$ from x_k?

2.2. What is a linesearch method for minimizing $f(x)$? Explain two ways in which a simple-minded linesearch might fail. Suggest how a backtracking Armijo strategy avoids these defects.

2.3. What are the three possible outcomes of the generic backtracking Armijo linesearch method?

2.4. Given these outcomes, show that the steepest-descent method applied to a bounded objective function must converge to a first-order critical point of $f(x)$.

2.5. Give two reasons why the steepest-descentr method may be poor in practice.

2.6. What restrictions are required to ensure that the search direction s_k computed from $B_k p_k = -\nabla f(x_k)$ will result in the backtracking Armijo linesearch finding a first-order critical point of $f(x)$?

2.7. Given conditions under which the choice $B_k = \nabla_{xx} f(x_k)$ will lead to a rapidly converging method.

2.8. Explain two ways in which an indefinite Hessian $\nabla_{xx} f(x_k)$ may be altered to ensure that the lineesearch method will succeed, 2.9. What is the secant condition? Show that the symmetric rank-one method satisfies the secant condition.

2.10. Find conditions which define the minimizer of the quadratic function $q(p) = p^T g + \frac{1}{2} p^T H p$ when $p = D p_d$ for a given matrix D.

2.11. What does it mean for a set of vectors $\{d^i\}$ to be B-conjugate?

2.12. What is important about the set of vectors generated by the conjugate-gradient method when approximately minimizing $q(p)$? What is the first such direction?

Part 3: Trust-region methods for unconstrained minimization

3.1. What is the solution of the trust-region problem when the model is a first-order (linear) Taylor's approximation to $f(x)$ around x at the point $x + s$?

3.2. Give an expression for the Cauchy point for the model derived from a second-order (quadratic) Taylor's approximation to $f(x)$ around x at the point $x + s$.

3.3. Derive a lower bound for the decrease in this quadratic model at the Cauchy point.

3.4. Argue (roughly) why a trust-region method must be able to make progress at a non-critical point by reducing the trust-region radius.

3.5. State necessary and sufficient conditions for the global minimizer of a quadratic function within an ℓ_2-norm trust region.

3.6. Explain how these optimality conditions might be used as the basis for a method for solving the ℓ_2-norm trust-region subproblem.

3.7. If the model is a second-order Taylor's approximation, what is the solution to the trust-region subproblem when the radius becomes very large?

3.8. What is the "hard" case? Suggest how the solution to the ℓ_2-norm trust-region subproblem.may be found in this case.

3.9. What is the secular equation? Derive Newton's method for finding a root of this equation.

3.10. Explain how the conjugate-gradient method may be modified to find an approximate solution to the ℓ_2-norm trust-region subproblem. How good is this solution when the quadratic model is convex?

Part 4: Active-set methods for linearly-constrained minimization

4.1. What do we mean by a convex quadratic function? What about a strictly convex quadratic one?

4.2. Show that there is only one local minimum of a convex quadratic. Show that the minimizer is unique if the function is strictly convex.

4.3. Give an example of a non-convex quadratic program in n unknowns that has 2^n distinct local minimizers.

4.4. State first-order necessary optimality conditions for a strictly convex quadratic program. Explain very generally how active-set and interior point methods use thees optimality conditions in different ways.

4.5. Give a method for finding a basis for the null-space of the m by n matrix A where $m \leq n$.

4.6. Give four possible outcomes of trying to minimize a quadratic function $q(x)$ over an affine set $Ax = 0$. Say what we mean by a direction of negative curvature. Give a necessary and sufficient condition for $q(x)$ to have such direction.

4.7. Suppose that $q(x)$ is strictly convex. Explain the differences between full-space, range-space and null-space methods for minimizing $q(x)$ over $Ax = 0$.

4.8. In a primal active-set method, if the current iterate is feasible and minimizes the objective over the current working set of active constraints, when does this iterate solve the problem over the complete set of constraints?

4.9. Give two ways of finding an initial feasible point for use in an active-set quadratic programming method.

Part 5: Penalty and augmented Lagrangian methods for constrained minimization

5.1. Explain what properties a merit function should have for general constrained optimization.

5.2. What is the quadratic penalty function for equality-constrained optimization? What is the gradient and Hessian of this function?

5.3. Show that first-order critical points of the quadratic penalty function approaches those of the related equality-constrained optimization problem as the penalty parameter approaches zero; state any assumptions used. What might happen if your assumption does not hold?

5.4. How do the eigenvalues of the Hessian matrix of the quadratic penalty function behave as the penalty parameter approaches zero?

5.5. Show how that the ill-conditioning suggested by the previous result does not mean that the Newton direction cannot be computed accurately.

5.6. What is the augmented Lagrangian function for equality-constrained optimization? What is the gradient and Hessian of this function?

5.7. Indicate how the augmented Lagrangian function may be manipulated so as to prevent the need for the penalty parameter to shrink to zero

Part 6: Interior-point methods for constrained minimization

6.1. What is the logarithmic barrier function for inequality-constrained optimization? What is the gradient and Hessian of this function?

6.2. Suppose that all of the inequality constraints are active, and have independent gradients, at limit points of the minimizer of the logarithmic barrier function as the barrier parameter approaches zero. Show that the resulting point is a critical point for the related inequality-constrained optimization problem.

6.3. How do the eigenvalues of the Hessian matrix of the logarithmic barrier function behave if the iterates approach the boundary of the feasible region? Is this a fatal flaw?

6.4. How, approximately, do the Newton equations for the logarithmic barrier function behave close to its minimizers as the barrier parameter shrinks to zero?

6.5. Argue that the primal Newton step is not a direction.

6.6. What are primal-dual path following methods? What is the central path?

6.7. Argue that the primal-dual Newton step is a good direction.

6.8. Suggest how a feasible point may be found for a general system of inequalities.

Part 7: SQP methods for constrained minimization

7.1. Derive Newton's method for satisfying the first-order necessary optimality conditions for an equality-constrained optimization problem.

7.2. How do these relate to minimizing a quadratic approximation to the objective function subject to linear approximations to the constraints?

7.3. What, generically, is a sequential quadratic programming method? What difficulties might arise in its pure form?

7.4. Show that the SQP direction is a descent direction for the quadratic penalty function if the penalty parameter is sufficiently small. What assumption are you making about the Hessian of the quadratic approximation in this case?

7.5. What is a dual norm? What is the dual norm of the ℓ_2 norm?

7.6. Give an example of a non-differentiable penalty function. Show that the SQP direction is a descent direction for your function; state any assumptions you make.

7.7. What is the Maratos effect? How can we try to overcome it?

7.8. Why might the obvious trust-region SQP method fail to work?

7.9. Explain the idea behind the Sℓ_1QP method.

7.10. Show that an ℓ_∞-norm trust-region ℓ_1QP subproblem may be written as an ordinary quadratic program.

7.11. What is a composite-step trust-region SQP method? Give an example.

7.12. What are the key ideas behind filter methods?