AMSC 607 / CMSC 764 Advanced Numerical Optimization
Fall 2006
Homework 2
Partial Solution

---

1. (10) Consider the Huang family of quasi-Newton formulas:

$$B^{(k+1)} = B^{(k)} - \frac{B^{(k)} s^{(k)} s^{(k)T} B^{(k)}}{s^{(k)T} B^{(k)} s^{(k)}} + \frac{y^{(k)} y^{(k)T}}{y^{(k)T} s^{(k)}} + \phi(s^{(k)T} B^{(k)} s^{(k)}) u^{(k)} u^{(k)T}$$

where ϕ is a scalar between 0 and 1 and

$$\begin{aligned} s^{(k)} &= x^{(k+1)} - x^{(k)}\,, \\ y^{(k)} &= g^{(k+1)} - g^{(k)}\,, \\ u^{(k)} &= \frac{y^{(k)}}{y^{(k)T} s^{(k)}} - \frac{B^{(k)} s^{(k)}}{s^{(k)T} B^{(k)} s^{(k)}}. \end{aligned}$$

(a) Show that this formula satisfies the secant condition.

(b) Show that if $B^{(k)}$ is positive definite and if $y^{(k)T} s^{(k)} > 0$, then $B^{(k+1)}$ is positive definite.

Answer:

(a) The secant condition says $B^{(k+1)} s^{(k)} = y^{(k)}$, and this can be checked by direct computation.

(b) We need to verify that $z^T B^{(k+1)} z > 0$ for all nonzero z. For the last term in $B^{(k+1)}$,

$$z^T u^{(k)} u^{(k)T} z = (z^T u^{(k)})^2 \geq 0,$$

and for the next to last,

$$z^T \frac{y^{(k)} y^{(k)T}}{y^{(k)T} s^{(k)}} z = \frac{(z^T y^{(k)})^2}{y^{(k)T} s^{(k)}} \geq 0$$

by the assumption. So we will just work with the first 2 terms, using a Cholesky factorization $B^{(k)} = LL^T$. Define $\bar{z} = L^T z$ and $\bar{s} = L^T s^{(k)}$. Then

$$\begin{aligned} z^T B^{(k)} z - z^T \frac{B^{(k)} s^{(k)} s^{(k)T} B^{(k)}}{s^{(k)T} B^{(k)} s^{(k)}} z &= \bar{z}^T \bar{z} - \frac{\bar{z}^T \bar{s} \bar{s}^T \bar{z}}{\bar{s}^T \bar{s}} \\ &= \frac{\bar{z}^T \bar{z} \bar{s}^T \bar{s} - \bar{z}^T \bar{s} \bar{s}^T \bar{z}}{\bar{s}^T \bar{s}} \\ &= \frac{\|\bar{z}\|^2 \|\bar{s}\|^2 - (\bar{z}^T \bar{s})^2}{\bar{s}^T \bar{s}}. \end{aligned}$$

The numerator is nonnegative (by the Cauchy-Schwarz inequality) since $\bar{z}^T \bar{s} = \|\bar{z}\| \|\bar{s}\| \cos\theta$ where θ is the angle between the two vectors. (In fact, it is positive unless $\bar{z}$ points in the direction $\pm\bar{s}$.)

Therefore, we know that $z^T B^{(k+1)} z \geq 0$ for all z. Now, if $\bar{z}$ points in the direction $\pm\bar{s}$, then the next to last term is strictly positive, so we can conclude that $z^T B^{(k+1)} z > 0$ for all nonzero z.

---

2. Consider the trust region method obtained by using $g(x)$ and $H(x)$ to form a quadratic model

$$f(x+p) \approx q(p) = f(x) + p^T g + \frac{1}{2} p^T H p$$

and then computing the search direction p from

$$\min_{\|p\|_2 \leq h} q(p)\,.$$

2(a) (5) Show that if H is positive definite, then the direction p is downhill.

2(b) (5) Suppose that the solution p to the minimization problem satisfies $\|p\|_2 < h$. In this case, give a formula for p.

Answer:

If the constraint $p^T p = h^2$ is active, then the solution is obtained from setting the derivative of the Lagrangian

$$L(x,\lambda) = f(x) + p^T g + \frac{1}{2} p^T H p + \frac{1}{2}\lambda(p^t p - h^2)$$

to zero:

$$\begin{aligned} g + Hp + \lambda p &= 0\,, \\ p^T p - h^2 &= 0\,. \end{aligned}$$

This gives

$$(H + \lambda I)p = -g\,.$$

The parameter λ is positive, since $\|p\|$ is a monotonically increasing function of λ.

If the constraint is inactive, then the solution is found by setting the derivative of q to zero, yielding the Newton direction.

$$Hp = -g$$

For $\lambda \geq 0$,

$$p^T (H + \lambda I)p = -p^T g,$$

and the left hand side is positive since H and $H + \lambda I$ are positive definite. Therefore, the direction is downhill.

---

3. (10) Consider the function

$$f(x, y) = \sin(x/y) - x^2 y.$$

Evaluate

$$\frac{\partial f}{\partial x} \text{ and } \frac{\partial f}{\partial y}$$

using the forward (bottom-up) mode of automatic differentiation.

Answer: See the posted notes.

---

4. (10) Consider the function

$$f(x, y) = \sin(x/y) - x^2 y.$$

Evaluate

$$\frac{\partial f}{\partial x} \text{ and } \frac{\partial f}{\partial y}$$

using the backward (top-down) mode of automatic differentiation.

Answer: See the posted notes.

---

5. (5) Consider the problem

$$\min_{x \geq 0} f(x)$$

where $x \in \mathcal{R}^5$. Suppose we have a current guess

$$\hat{x} = [0, 5, 3, 0, 2]^T$$

for the solution and that we have computed the corresponding Lagrange multipliers

$$\hat{\lambda} = [1, -2, 3, -4, 5]^T .$$

Write the first order necessary conditions for optimality. For each condition, indicate whether the point $\hat{x}, \hat{\lambda}$ satisfies, violates, or possibly satisfies it.

Answer:

1. $\hat{x}$ is feasible since $\hat{x} \geq 0$.
2. We don't know whether $\hat{\lambda} = g(\hat{x})$.
3. The condition $\hat{\lambda} \geq 0$ is violated for variables $2, 4$.
4. The complementarity conditions $\hat{\lambda}_i \hat{x}_i = 0$ are violated for $i = 2, 3, 5$.

---

6. (10) Consider the linear programming problem

$$\begin{aligned} \min_x \quad & c^T x \\ Ax &= b \\ x &\geq 0 \end{aligned}$$

where A is an $m \times n$ matrix with $m < n$.

Show that if there exists a solution to this problem, then there is a solution x_{opt} that has at most m nonzero components.

Answer: First, consider this example:

$$\min_x x_1 + x_2$$

subject to

$$\begin{aligned} x_1 + x_2 &= 3 \\ x &\geq 0 \end{aligned}$$

The set of solutions to this problem is $\{[\alpha, (3-\alpha)]^T : 0 \leq \alpha \leq 3\}$. So clearly there are solutions for which all components of x are nonzero.

So let's answer the question. Suppose x solves our problem, and at least $m+1$ elements of x are nonzero.

- The optimality conditions say that x is feasible, $[I, A^T]\lambda = c$, and $\lambda^T x = 0$. Note that there are $n+m$ Lagrange multipliers λ.
- Since at least $m+1$ elements of x are nonzero, at least $m+1$ components of λ must be zero (by complementarity), so we have expressed c as a linear combination of at most $n-1$ constraint gradients.
- Let $\hat{A}^T$ be the matrix whose columns are these constraint gradients. We have found a vector $\hat{\lambda}$ satisfying

  $$\hat{A}^T \hat{\lambda} = c$$

  or

  $$\begin{bmatrix} \hat{A}^T & c \end{bmatrix} \begin{bmatrix} \hat{\lambda} \\ -1 \end{bmatrix} = 0.$$

  Therefore, the matrix is rank deficient.
- Therefore, there is a vector z for which

  $$\begin{bmatrix} \hat{A} \\ c^T \end{bmatrix} z = 0.$$

  Moving in this direction keeps all of the active constraints satisfied and does not change the objective function value.

- Change x to $x + \alpha z$, where α is the largest step we can take without making some component negative; then $x + \alpha z$ has at least one more zero than x has.

Repeat this process until at most m components of x are nonzero.

---

7. (15) Write a MATLAB function to minimize a given function f subject to linear equality constraints $Ax = b$. Your function should use the feasible direction formulation, using the QR factors to get a basis for the feasible directions, along with MATLAB's `fminunc`. It should be written to find a local minimizer for arbitrary functions f and arbitrary numbers of variables and constraints. The user should be given the option of providing a starting point.

Test your function on the problem

$$\min_x x_1^2 x_2^3 + 4x_1^2 x_3^2 + x_2^4 x_3^2 + 3x_1 x_2 + 4x_2 x_3 + 5x_1 x_3 + x_1 + x_3$$

subject to the constraint

$$x_1 + x_2 + x_3 = 3,$$

starting with the point $x = [-1, 5, -1]^T$. Print out the sequence of iterates and document your function well.

Answer: Geping Liu's solution is posted. Note that if the initial point is specified, it should be checked for feasibility, and if it is not specified, it should be generated.