MATH 519/619 Optimization — Homework set #2

Winter quarter 2006 — Miguel Á. Carreira-Perpiñán

This set covers chapters 5, 6, 8 and 10 of the book *Numerical Optimization* by Nocedal and Wright.

From the book, the following exercises: 5.1, 5.2, 5.3, 5.6, 5.8, 5.11, 6.1, 8.1, 8.3, 8.6. In addition, the exercises below. The Matlab programming exercises are 5.1, 5.8, 6.1 and **2**. For exercise 5.8, try 2 different matrices $\mathbf{A}$ generated as follows:

```
% Matrix 1: uniform spectrum
n = 100; l = 1:n; U = gallery('orthog',n); A = U*diag(l)*U';
% Matrix 2: clustered spectrum; try s = 5, 1, 0.001, 0
n = 100; n1 = floor(n/4);
l = [ones(1,n1) n/5*ones(1,n1) n*ones(1,n-2*n1)]; rand('state',314); l = l + s*rand(1,n);
U = gallery('orthog',n); A = U*diag(l)*U';
```

and compare your results with figs. 5.4–5.5. For exercise 6.1, use the `convseq` function you wrote in exercise hw1–**4** to estimate the convergence rate. Hint for exercise 8.1(a): on a convex domain $\mathcal{D}$, f is strongly convex $\Leftrightarrow \forall \mathbf{x}, \mathbf{y} \in \mathcal{D}: \; f(\mathbf{y}) > f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x})$ (this is called *first-order convexity condition*).

Exercise 1. Let $\mathbf{A}$ be a real, symmetric, positive definite, $n \times n$ matrix with eigenvectors $\mathbf{U} = (\mathbf{u}_1, \dots, \mathbf{u}_n)$ and associated eigenvalues $\boldsymbol{\Lambda} = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ (in matrix form). Show that: (i) If $\mathbf{R} = \mathbf{U}\boldsymbol{\Lambda}^{1/2}\mathbf{Q}$, where $\mathbf{Q}$ is any orthogonal matrix, then $\mathbf{A} = \mathbf{R}\mathbf{R}^T$. (ii) If $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ are nonzero orthogonal vectors then $\{\mathbf{p}_1, \dots, \mathbf{p}_n\}$ where $\mathbf{p}_i = (\mathbf{R}^{-1})^T\mathbf{v}_i$ are conjugate w.r.t. $\mathbf{A}$ (this shows there is an infinite number of conjugate direction sets). (iii) $\mathbf{p}_i = \mathbf{u}_i$ are conjugate w.r.t. $\mathbf{A}$ (this corresponds to the change of variables $\hat{\mathbf{x}} = \mathbf{S}^{-1}\mathbf{x}$ with $\mathbf{S} = \mathbf{U}$). Hint: use the spectral theorem.

Exercise 2. Implement algorithm 8.1 (BFGS method) in Matlab with $\mathbf{H}_0 = \mathbf{I}$ and backtracking line search (with initial step length 1). Apply it to the Rosenbrock function (2.23) from two initial points $\mathbf{x}_0 = (1.2, 1.2)$ and $\mathbf{x}_0 = (-1.2, 1)$ (cf. exercise 3.1). Tabulate $\|\mathbf{x}_k - \mathbf{x}^*\|_2$, $\|\nabla f(\mathbf{x}_k)\|_2$ and $\left\|\mathbf{B}_k - \nabla^2 f(\mathbf{x}_k)\right\|_F$ where $\|\cdot\|_2$ is the Euclidean norm (for vectors) and $\|\cdot\|_F$ the Frobenius norm (for matrices). Use your `convseq` function to estimate the convergence rate.

Exercise 3. Verify that, in 1D, all 3 quasi-Newton methods BFGS, DFP and SR1 are equivalent to the secant method, where $B_{k+1} = \frac{f'_{k+1} - f'_k}{x_{k+1} - x_k}$ independent of B_k.

Exercise 4. Consider Newton's method (algorithm 6.2 in p. 141 of the book) with a Hessian modification $\mathbf{B}_k = \nabla^2 f(\mathbf{x}_k) + \lambda_k \mathbf{I}$, so that the search direction is $\mathbf{p}_k = -\mathbf{B}_k^{-1}\nabla f(\mathbf{x}_k)$, and has as extreme cases the pure Newton step for $\lambda = 0$ and the steepest descent direction for $\lambda \to \infty$. Assume that $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x}$ and $\mathbf{x}_k = (1, -3)$. Plot the contours of f, the gradient at $\mathbf{x}_k$ and the search direction $\mathbf{p}_k$ as a function of $\lambda \geq 0$ for the following three cases: $\mathbf{A} = \left(\begin{smallmatrix} 4 & 0 \\ 0 & -1 \end{smallmatrix}\right)$; $\mathbf{A} = \left(\begin{smallmatrix} 4 & 0 \\ 0 & 1 \end{smallmatrix}\right)$; $\mathbf{A} = \left(\begin{smallmatrix} -4 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. What is the relation with theorem 4.3 (p. 78 in the book) and with exercise 4.1?