AMSC/CMSC 666

NUMERICAL ANALYSIS I

HOMEWORK # 4 (Pbs 1-3 due Th 11/14, Pbs 4-7 due Tu 11/26)

1 (15 pts). *Derivation and properties of BFGS.* Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$ be a C^1 function. Let $B_k = L_k L_k^T$ be the Cholesky decomposition of the SPD matrix $B_k \in \mathbb{R}^{n\times n}$. Recall that one step of the quasi-Newton method reads

$$s_k = x_{k+1} - x_k = -B_k^{-1}\nabla f(x_k), \quad y_k = \nabla f(x_{k+1}) - \nabla f(x_k). \tag{1}$$

To construct the next secant matrix $B_{k+1} \in \mathbb{R}^{n\times n}$ we enforce the form $B_{k+1} = J_{k+1}J_{k+1}^T$ for a suitable nonsingular matrix $J_{k+1} \in \mathbb{R}^{n\times n}$ and vector $v_k \in \mathbb{R}^n$ such that

$$J_{k+1}v_k = y_k, \quad J_{k+1}^T s_k = v_k$$

(a) Let J_{k+1} be the Broyden update of L_k, namely

$$J_{k+1} = L_k + \frac{(y_k - L_k v_k)v_k^T}{v_k^T v_k},$$

and assume that $y_k^T s_k > 0$. The latter is guaranteed by the Wolfe conditions. Show that v_k must be parallel to $L_k^T s_k$ and is thus given by

$$v_k = \Big(\frac{y_k^T s_k}{s_k^T B_k s_k}\Big)^{1/2} L_k^T s_k.$$

(b) Show that B_{k+1} is given by

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k}. \tag{2}$$

(c) Show that B_{k+1} is SPD. To this end, prove that either J_{k+1} is nonsingular or $z^T B_{k+1} z > 0$ for all $0 \neq z \in \mathbb{R}^n$.

(d) Perform an operation count for one step. Note that the matrix-vector computation of $B_{k+1}^{-1}z$ can be performed using the Sherman-Morrison-Woodbury formula for rank-two updates, starting from the Cholesky decomposition of B_0.

2 (15 pts). *Scale invariance.* Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$ be a C^2 function. Show that the Newton method $x_{k+1} = x_k - D^2 f(x_k)\nabla f(x_k)$ and the BFGS quasi-Newton method (1) and (2) are scale invariant (or equivalently the iterates are independent of the units being used). Given a nonsingular matrix $S \in \mathbb{R}^{n\times n}$ and a vector $x \in \mathbb{R}^n$, consider the change of variables $x = S\tilde{x} + z$ and $\tilde{f}(\tilde{x}) = f(x)$, the initial conditions $x_0 = S\tilde{x}_0 + z$, and proceed as follows assuming unit step length $\lambda_k = 1$.

(a) Show that

$$\nabla \tilde{f}(\tilde{x}) = S^T \nabla f(x), \quad D^2\tilde{f}(\tilde{x}) = S^T D^2 f(x) S,$$

and that the Newton iterates obey the relation $x_k = S\tilde{x}_k + z$.

(b) Given an approximate Hessian B_0 of $D^2 f(x_0)$, let $\widetilde{B}_0 = S^T B_0 S$. Show that the BFGS iterates obey the relation $x_k = S\tilde{x}_k + z$.

3 (20 pts). *Performance of minimization methods.* Consider the Rosenbrock function in $\mathbb{R}^2$

$$f(x, y) = 100(y - x^2)^2 + (1 - x)^2. \tag{3}$$

(a) Plot the level sets of this function using the comamnd `contour`. Show that $(1, 1)$ is the only local minimizer and that the Hessian $D^2 f(1, 1)$ is positive definite.

(b) Program the steepest descent, Newton and BFGS methods using backtracking line search that imposes the Armijo-Wolfe conditions.

(c) Use these methods to minimize (3). First start with the initial guess $(1.2, 1.2)$ and then with the more difficult one $(1.2, 1)$. Set the initial step length $\lambda_0 = 1$ and plot the step length λ_k versus k for each of the methods. Plot the error $\|(x_k, y_k) - (1, 1)\|_2$ versus k in the logaritmic alone the vertical axis for each method. Do you observe superlinear convergence? Compare the performance of the methods.

4 (10 pts). *Gronwall inequality.* Let $\phi_n, \eta_n \geq 0$ and $0 \leq \chi_n < 1$ satisfy the recurrence inequality:

$$\phi_n \leq \eta_n + \sum_{k=1}^{n} \chi_k \phi_k \qquad (n \geq 1). \tag{4}$$

(a) Set $R_n = \sum_{k=1}^{n} \chi_k \phi_k$ and write (4) as $R_n - R_{n-1} - \chi_n R_n \leq \chi_n \eta_n$ for $n \geq 1$ with $R_0 = 0$. Multiply by the integrating factor $\prod_{j=1}^{n-1}(1-\chi_j)$, defined to be $=1$ for $n=1$, to derive the inequality

$$\phi_n \leq \eta_n + \sum_{k=1}^{n} \Big(\chi_k \eta_k \prod_{j=k}^{n} (1-\chi_j)^{-1} \Big).$$

(b) If η_n is non-decreasing, show the simplified expression

$$\phi_n \leq \eta_n \prod_{j=1}^{n} (1-\chi_j)^{-1}. \tag{5}$$

(c) Let $\chi_n = h_n L$ where $h_n = t_n - t_{n-1}$ is a variable stepsize ($n \geq 1$) and $T = \sum_{n-1}^{N} h_n$. Prove that (5) becomes the following more familiar expression provided $\chi_n \leq 1/2$:

$$\phi_n \leq \eta_n e^{\sum_{j=1}^{n} \chi_j} = \eta_n e^{2LT}. \tag{6}$$

5 (15 pts). *Trapezoidal method.* Consider the method with variable stepsize h_n

$$y_n = y_{n-1} + \frac{h_n}{2} \Big[f(t_{n-1}, y_{n-1}) + f(t_n, y_n) \Big]. \tag{7}$$

(a) Find a condition on h_n such that the nonlinear problem (7) admits a unique solution y_n.

(b) Find an expression for the truncation error and prove a global error estimate (Hint: use Taylor's formula and (6)).

(c) Remove the exponential constant in (b) provided f is dissipative (or equivalently $-f$ is monotone).

6 (10 pts). *Predictor-Corrector method.* Consider the following method as a way to approximate (7):

$$\begin{aligned} y_n^{(0)} &= y_{n-1} + h_n f(t_{n-1}, y_{n-1}) \\ y_n^{(j)} &= y_{n-1} + \frac{h_n}{2} \Big(f(t_{n-1}, y_{n-1}) + f(t_n, y_n^{(j-1)}) \Big), \qquad j \geq 1. \end{aligned}$$

(a) Determine under which condition(s) this fixed point iteration converges, that is $y_n^{(j)} \to y_n$ as $j \to \infty$ where y_n is the solution of (7).
(b) Show that $\|y_n - y_n^{(j)}\| = O(h_n^{2+j})$. Conclude that one iteration is sufficient to preserve the truncation error of (7). The resulting scheme for $j=1$ is a 2nd order Runge-Kutta Method with variable stepsize, the so-called *Heun's method.*

7 (15 pts). *Absolute stability.* Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be a symmetric and positive definite matrix. For $\mathbf{F} : \mathbb{R} \to \mathbb{R}^n$ being given, consider the linear system of n ODEs

$$\mathbf{y}' + \mathbf{A}\mathbf{y} = \mathbf{F}. \tag{8}$$

The system (8) arises, for instance, from a space discretization via finite differences or finite elements of the Heat Equation $u_t - \Delta u = f$. The resulting matrix $\mathbf{A}$ exhibits very disparate eigenvalues (*stiff* system).

(a) Determine the magnification factor for the Forward Euler (FE), Backward Euler (BE), Trapezoidal Methods (TM), and Heun's method (HM). Such a factor is expressed in terms of the spectral radius of a suitable matrix that is to be found.

(b) Discuss the resulting stability constraints and relate to absolute stability.

(c) Explain why BE and TM, which reduces to the Crank-Nicholson method for linear ODEs, are always preferred over FE and HM for time stepping (8).