Suggestion Answer for Homework 5, CS421

November 29, 1999

1. There is a typo in the problem. We need $u = x - sgn(x_1)\|x\|e_1$. Since

$$u^T x = x^T x - sgn(x_1)\|x\|(e_1^T x) = \|x\|^2 - |x_1|\|x\|$$

$$u^T u = x^T x + \|x\|^2 - 2(e_1^T x)sgn(x_1)\|x\| = 2(\|x\|^2 - |x_1|\|x\|,$$

we have

$$Qx = x - u = sgn(x_1)\|x\|e_1.$$

2.
The details is on *Matrix computations, Van Loan, P210 - P225.*
Assume A is m-by-n matrix. Suppose function *house(v)* returns householder vector. This function needs $3n$ flops. The algorithm (*Matrix computations , P224*) for QR factorization requires $2n^2(m - n/3)$ flops. Computing Q explicitly requires $4(m^2n - mn^2 + n^3/3)$ flops.

3 True.
The dominant work to solve a square linear system via Householder transformations and a backsolve is QR factorization. From problem 2, it requires $4n^3/3$ flops.
The dominant work for LU factorization requires $2n^3/3$ flops.

4.
Assume $A = Q\ R$ where Q is orthogonal matrix and R is upper triangular matrix. Denote $\tilde{A}$ to be the matrix formed by row permutations of A. Then there exists an elementary matrix E such that $\tilde{A} = E\ A$.
$\Longrightarrow$

$$\tilde{A} = E\ Q\ R = \tilde{Q}\ R$$

where $\tilde{Q} = E\ Q$. Since $\tilde{Q}$ is also orthogonal matrix, it is QR factorization of $\tilde{A}$. Thus there is no change of sparsity of R.

5.
Assume there are k groups formed by coloring of the intersection graph. Let $\{d_1, \cdots, d_k\}$ be the vector corresponding to each group. Let λ_i be the number of nonzero entries in the

Jocobian matrix corresponding to each group. Let $\tilde{J}$ be the vector whose elements are the nonzero entries of J. Assume the elements of $\tilde{J}$ are sorted by the order of group. Then

$$diag[\alpha I_{\lambda_1}, \cdots, \alpha I_{\lambda_k}]\tilde{J} = [F(x+\alpha d_1) - F(x); \cdots; F(x+\alpha d_k) - F(x)]$$

where I_{λ_i} is identity matrix with size λ_i

6(a).
Let us take QR factorization of A^T. Assume

$$A^T = [Q_1\ Q_2]\begin{bmatrix} R \\ 0 \end{bmatrix}$$

For any column vector z in Q_2, we claim that $Az = 0$. This is because

$$Az = [R^T\ 0]\begin{bmatrix} Q_1^T z \\ Q_2^T z \end{bmatrix} = [R^T\ 0]\begin{bmatrix} 0 \\ Q_2^T z \end{bmatrix} = 0.$$

Thus

$$Z = Q_2$$

(b) Assume $Ax_0 = b$. Then for any solution x we have the form

$$x = x_0 + Zy$$

where y is $n-m$-vector. Then

$$q(x) = \frac{1}{2}y^T Z^T H_0 Z y + (c^T Z + x_0^T H Z)y + (c^T x_0 + \frac{1}{2}x_0^T H x_0)$$

Denote $\phi(y) = q(x)$. The equivalent problem is $\min\{\phi(y); y \in R^{n-m}\}$. The necessary condition for the existence of the local minimum solution is that there exists y^* such that $Z^T H Z y^* = -Z^T H x_0 - Z^T c$. If H is SPD, $Z^T H Z$ is also SPD. Thus there is an unique solution for the problem. If H is positive semi-definite and the necessary condition is satisfied, there exists at least one solution for the problem. If $Z^T H Z$ is full rank, the solution is unique, otherwise there exists infinitely many solutions.

7.
There is a demo code in Optimization Toolbox of MATLAB. The command is 'bandemo'. The initial starting point of the demo is $(-2, 2)$.