The University of Connecticut **Fall 2009**
Dept. of ESE **KRP**

Problem Set # 5

(Due Date: October 13, 2009)

1. Recall the Gram-Schmidt method for generating the Q-conjugate directions. Suppose that $\{\underline{p}_i\}_{i=0}^{n-1}$ are generated as moments of Q, that is, suppose $\underline{p}_k = Q^k \underline{p}_0, k = 1, 2, .., n-1.$ This is the so-called Krylov subspace. Show that the corresponding $\underline{d}_k$'s can then be generated by a (three-term) recursion formula where $\underline{d}_{k+1}$ is defined in terms of $Q\underline{d}_k$, $\underline{d}_k$ and $\underline{d}_{k-1}$. (Hint: This recursion is related to Lanczos tri-diagonalization of symmetric matrices.)

2. Suppose that the $\{\underline{p}_i\}_{i=0}^{n-1}$ in Problem 1 are taken as $\underline{p}_i = \underline{e}_i$, where $\underline{e}_i$ is the i^{th} unit coordinate vector and the $\underline{d}_k$'s are constructed accordingly. Show that using $\underline{d}_k$'s in a conjugate direction method to minimize $f(\underline{x}) = \frac{1}{2}\underline{x}^T Q \underline{x} - \underline{b}^T \underline{x}$ is equivalent to the application of Gaussian elimination to solve $Q\underline{x} = \underline{b}$.

3. Show that in the purely quadratic form of the conjugate gradient method $\underline{d}_k^T Q \underline{d}_k = -\underline{d}_k^T Q \underline{g}_k$. Using this show that to obtain $\underline{x}_{k+1}$ from $\underline{x}_k$, it is necessary to use Q only to evaluate $\underline{g}_k$ and $Q\,\underline{g}_k$.

4. Show that in the quadratic problem $Q\,\underline{g}_k$ can be evaluated by taking a unit step from $\underline{x}_k$ in the direction of the negative gradient and evaluating the gradient there. Specifically, if $\underline{y}_k = \underline{x}_k - \underline{g}_k$ and $\underline{p}_k = \nabla \underline{f}(\underline{y}_k)$, then $Q\underline{g}_k = \underline{g}_k - \underline{p}_k$. Combine this result with the result of Problem 3 to derive a conjugate gradient method for general problems which does not require the knowledge of the Hessian or a line search (but it need not have global convergence properties. For global convergence, you need line search!).