Math 408 Homework Set 5

(1) Let $Q \in \mathbb{R}^{n\times n}$ be symmetric and positive definite, and let $d_1, d_2, \ldots, d_n \in \mathbb{R}^n$ be a basis of Q-conjugate vectors. Show that

$$Q^{-1} = \sum_{i=1}^{n} \frac{d_i d_i^T}{d_i^T Q d_i} \ .$$

(2) Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$, and suppose that $\mathrm{Nul}(A) = \{0\}$. Consider the function

$$f(x) := \frac{1}{2}\|Ax - b\|_2^2 \ .$$

Show how to apply the CG algorithm to minimize f on $\mathbb{R}^n$ under the assumption that A is not available but Ax and $A^T y$ can be obtained for arbitrary vectors $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$, respectively.

(3) Apply the conjugate gradient algorithm initialized at $x^0 = (0,0,0)^T$ to solve the problem $\min_{x\in\mathbb{R}^3} \frac{1}{2}x^T Hx + g^T x$, where

$$H = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix} \quad \text{and} \quad g = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} .$$

(4) Let $f(x) := \frac{1}{2}x^T Hx + g^T x$, where $H \in \mathcal{S}^n_{++}$ and $g \in \mathbb{R}^n$. Let $\{x^0, \ldots, x^n\}$ be the iterates of the conjugate gradient algorithm applied to f initiated at $x^0 = 0$.

- (a) Show that $f(x^k) - f(x^{k+1}) = \frac{1}{2}\frac{\langle g^k, d^k\rangle)^2}{\langle Hd^k, d^k\rangle}$, which establishes that $f(x^k) > f(x^{k+1})$ whenever $d^k \neq 0$.
- (b) Show that the sequence $\{\beta_k\}$ generated by the CG-algorithm satisfies $\beta_k = \frac{\langle g^{k+1}, g^{k+1} - g^k\rangle}{\|g^k\|^2}$.

(5) Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$, and consider the function

$$f(x) := \frac{1}{2}\|Ax - b\|_2^2 \ .$$

Show that $\nabla f(x) = A^T(Ax - b)$ and $\nabla^2 f(x) = A^T A$.

(6) Let $f(x) := \frac{1}{2}x^T Hx + g^T x$, where $H \in \mathbb{R}^{n\times n}$ and $g \in \mathbb{R}^n$.

- (a) Show that $\nabla f(x) = \frac{1}{2}(H + H^T)x + g$ so that if $H \in \mathcal{S}^n$, then $\nabla f(x) = Hx + g$.
- (b) Show that $\nabla^2 f(x) = \frac{1}{2}(H + H^T)$ so that if $H \in \mathcal{S}^n$, then $\nabla^2 f(x) = H$.

(7) Let $F : \mathbb{R}^n \to \mathbb{R}^m$ be a smooth function (has as many continuous derivatives as you want), and define $f(x) := \frac{1}{2}\|F(x)\|_2^2$.

- (a) Show that $\nabla f(x) = \nabla F(x)^T F(x)$.
- (b) Show that $\nabla^2 f(x) = \nabla F(x)^T \nabla F(x) + \sum_{j=1}^m F_j(x)\nabla^2 F_j(x)$, where $F_j : \mathbb{R}^n \to \mathbb{R}$ are the component functions of F, $j = 1, \ldots, m$.

(8) Let $f : \mathbb{R}^n \to \mathbb{R}$ be twice continuously differentiable, $V \in \mathbb{R}^{n\times k}$, and $\hat{x} \in \mathbb{R}^n$, where $k < n$. Consider the function given by $h(z) := f(\hat{x} + Vz)$.

- (a) Show that $\nabla h(z) = V^T \nabla f(\hat{x} + Vz)$.
- (b) Show that if $\bar{z}$ is a local solution to $\min_z h(z)$, then $\nabla f(\bar{z}) \perp \mathrm{Ran}\,(V)$.
- (c) Let S be a subspace of $\mathbb{R}^n$. Show that if $\bar{x} \in \hat{x} + S$ solves $\min\{f(x) \mid x \in \hat{x} + S\}$, then $\nabla f(\bar{x}) \perp S$.