1 More on Conjugate Directions (July 16)

1.1 Geometric Interpretation

Let $d_0, \ldots, d_{n-1}$ be a set of non-zero Q-orthogonal vectors in $\mathbb{R}^n$, where Q is symmetric and positive definite. Note that these vectors are linearly independent by a result from last lecture. Let B_k denote the subspace spanned by the first k vectors. We have an increasing sequence

$$B_0 \subsetneq B_1 \subsetneq \cdots \subsetneq B_n,$$

and $\dim(B_k) = k$.

Theorem 1.1. *The sequence $\{x_k\}_{k=0}^{\infty}$ generated from x_0 by the method of conjugate directions has the property that x_k minimizes $f(x) = \frac{1}{2}x^T Qx - b^T x$ on the affine subspace $x_0 + B_k$.*

Recall the function $q(x) = \frac{1}{2}(x - x_*)^T Q(x - x_*)$, which differs from $f(x)$ by a constant. They have the same minimizers.

If $Q = I$, then $q(x) = \frac{1}{2}|x - x_*|^2$. Then $x_k \in x_0 + B_k$ is the closest point in $x_0 + B_k$ to x_*, by the theorem.

Before proving the theorem, recall the following result about convex functions.

Lemma 1.2. *Let f be a C^1 convex function defined on a convex domain $\Omega \subseteq \mathbb{R}^n$. Suppose there is an $x_* \in \Omega$ such that $\nabla f(x_*) \cdot (y - x_*) \geq 0$ for all $y \in \Omega$. Then x_* is a global minimizer of f on Ω. The converse is obviously true.*

Geometrically, this means that if we move in any feasible direction from the point x_*, the function is increasing. Hence x_* is a local minimizer; convexity implies it is global. With this result in mind, we prove the theorem.

Proof. The affine subspace $\Omega = x_0 + B_k$ is convex. **(This proof could not be finished as attention had to be diverted from the lecture.)** □

Corollary 1.2.1. *x_n minimizes $f(x)$ on $\mathbb{R}^n$. That is, $x_n = x_*$; the method of conjugate directions for this function f terminates in at most n steps.*

When $Q = I$, then $q(x)$ is half the distance squared from x to x_*. What if $Q \neq I$. q is still a metric on $\mathbb{R}^n$. Thus x_k is the point "closest" to x_* on the affine subspace $x_0 + B_k$. **(These notes are incomplete.)**