in which the $g^{(k)}$, $k = 1, 2, \ldots$, are bounded away from zero when the Polak-Ribière formula is used (see [77]). In the study by Powell in [77], a global convergence analysis suggests that the Fletcher-Reeves formula for β_k is superior. Powell further suggests another formula for β_k:

$$\beta_k = \max\left[0, \frac{g^{(k+1)T}[g^{(k+1)} - g^{(k)}]}{g^{(k)T}g^{(k)}}\right].$$

For general results on the convergence of conjugate gradient methods, we refer the reader to [98].

10.2
10.5
10.7

EXERCISES

10.1 (Adopted from [64, Exercise 8.8(1)]) Let Q be a real symmetric positive definite $n \times n$ matrix. Given an arbitrary set of linearly independent vectors $\{p^{(0)}, \ldots, p^{(n-1)}\}$ in $\mathbb{R}^n$, the *Gram-Schmidt* procedure generates a set of vectors $\{d^{(0)}, \ldots, d^{(n-1)}\}$ as follows:

$$\begin{aligned} d^{(0)} &= p^{(0)} \\ d^{(k+1)} &= p^{(k+1)} - \sum_{i=0}^{k} \frac{p^{(k+1)T}Qd^{(i)}}{d^{(i)T}Qd^{(i)}} d^{(i)}. \end{aligned}$$

Show that the vectors $d^{(0)}, \ldots, d^{(n-1)}$ are Q-conjugate.

10.2 Let $f : \mathbb{R}^n \to \mathbb{R}$ be the quadratic function

$$f(x) = \frac{1}{2}x^TQx - x^Tb,$$

where $Q = Q^T > 0$. Given a set of directions $\{d^{(0)}, d^{(1)}, \ldots\} \subset \mathbb{R}^n$, consider the algorithm

$$x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)},$$

where α_k is the step size. Suppose that $g^{(k+1)T}d^{(i)} = 0$ for all $k = 0, \ldots, n-1$ and $i = 0, \ldots, k$, where $g^{(k+1)} = \nabla f(x^{(k+1)})$. Show that if $g^{(k)T}d^{(k)} \neq 0$ for all $k = 0, \ldots, n-1$, then $d^{(0)}, \ldots, d^{(n-1)}$ are Q-conjugate.

10.3 Let $f : \mathbb{R}^n \to \mathbb{R}$ be given by $f(x) = \frac{1}{2}x^TQx - x^Tb$, where $b \in \mathbb{R}^n$, and Q is a real symmetric positive definite $n \times n$ matrix. Show that in the conjugate gradient method for this f, $d^{(k)T}Qd^{(k)} = -d^{(k)T}Qg^{(k)}$.

10.4 Let Q be a real $n \times n$ symmetric matrix.

a. Show that there exists a Q-conjugate set $\{d^{(1)}, \ldots, d^{(n)}\}$ such that each $d^{(i)}$

Hint: Use the fact that for any real symmetric $n \times n$ matrix, there exists a set $\{v_1, \ldots, v_n\}$ of its eigenvectors such that $v_i^T v_j = 0$ for all $i, j = 1, \ldots, n$, $i \neq j$.

b. Suppose that Q is positive definite. Show that if $\{d^{(1)}, \ldots, d^{(n)}\}$ is a Q-conjugate set that is also orthogonal (i.e., $d^{(i)T}d^{(j)} = 0$ for all $i, j = 1, \ldots, n$, $i \neq j$), and $d^{(i)} \neq 0$, $i = 1, \ldots, n$, then each $d^{(i)}$, $i = 1, \ldots, n$, is an eigenvector of Q.

10.5 Consider the following algorithm for minimizing a function f:

$$x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)},$$

where $\alpha_k = \arg\min_\alpha f(x^{(k)} + \alpha d^{(k)})$. Let $g^{(k)} = \nabla f(x^{(k)})$ (as usual).

Suppose f is quadratic with Hessian Q. We choose $d^{(k+1)} = \gamma_k g^{(k+1)} + d^{(k)}$, and we wish the directions $d^{(k)}$ and $d^{(k+1)}$ to be Q-conjugate. Find a formula for γ_k in terms of $d^{(k)}$, $g^{(k+1)}$, and Q.

10.6 Consider the quadratic function $f : \mathbb{R}^n \to \mathbb{R}$ given by

$$f(x) = \frac{1}{2}x^TQx - x^Tb,$$

where $Q = Q^T > 0$. Let $D \in \mathbb{R}^{n \times r}$ be of rank r, and $x_0 \in \mathbb{R}^n$. Define the function $\phi : \mathbb{R}^r \to \mathbb{R}$ by

$$\phi(a) = f(x_0 + Da).$$

Show that ϕ is a quadratic function with a positive definite quadratic term.

10.7 Let $f(x)$, $x = [x_1, x_2]^T \in \mathbb{R}^2$, be given by

$$f(x) = \frac{5}{2}x_1^2 + \frac{1}{2}x_2^2 + 2x_1x_2 - 3x_1 - x_2.$$

a. Express $f(x)$ in the form of $f(x) = \frac{1}{2}x^TQx - x^Tb$.

b. Find the minimizer of f using the conjugate gradient algorithm. Use a starting point of $x^{(0)} = [0, 0]^T$.

c. Calculate the minimizer of f analytically from Q and b, and check it with your answer in part b.

10.8 Write a MATLAB routine to implement the conjugate gradient algorithm for general functions. Use the secant method for the line search (e.g., the MATLAB function of Exercise 7.9). Test the different formulas for β_k on Rosenbrock's function (see Exercise 9.3), with an initial condition $x^{(0)} = [-2, 2]^T$. For this exercise, reinitialize the update direction to the negative gradient every 6 iterations.