Caleb Logemann
MATH 565 Continuous Optimization
Homework 4

1. Let $\underline{\underline{A}} \in \mathbb{R}^{n\times n}$ be a symmetric positive definite matrix.

   (a) Show that the unit vectors $\underline{d}_1, \underline{d}_2, \dots, \underline{d}_n$ are $\underline{\underline{A}}$-conjugate vectors if and only if $\underline{\underline{D}}^T\underline{\underline{A}}\underline{\underline{D}} = \underline{\underline{A}}$ where $\underline{\underline{D}} = [\underline{d}_1, \underline{d}_2, \dots, \underline{d}_n]$.

   Proof. First assume that $\{\underline{d}_1, \underline{d}_2 \dots, \underline{d}_n\}$ are $\underline{\underline{A}}$-conjugate. Consider the matrix $\underline{\underline{D}} = [\underline{d}_1, \underline{d}_2, \dots, \underline{d}_n]$. Then the matrix multiplciation $\underline{\underline{D}}^T\underline{\underline{A}}\underline{\underline{D}}$ results in another matrix whose ij entry is $\underline{d}_i^T\underline{\underline{A}}\underline{d}_j$. If $i \neq j$, then $\underline{d}_i^T\underline{\underline{A}}\underline{d}_j = 0$ because $\underline{d}_i$ and $\underline{d}_j$ are $\underline{\underline{A}}$-conjugate. If on the other hand $i = j$, then $\underline{d}_i^T\underline{\underline{A}}\underline{d}_i = 0$ becuase $\underline{d}_i$ is a unit vector with respect to the $\underline{\underline{A}}$-norm. This shows that $\underline{\underline{D}}^T\underline{\underline{A}}\underline{\underline{D}} = \underline{\underline{I}}$, because entries on the diagonal are 1 and entries off the diagonal are 0.
   Now assume that $\underline{\underline{D}}^T\underline{\underline{A}}\underline{\underline{D}} = I$, and consider the set $\{\underline{d}_1, \underline{d}_2, \dots, \underline{d}_n\}$. Since $\underline{\underline{D}}^T\underline{\underline{A}}\underline{\underline{D}} = I$, this implies that $\underline{d}_i^T\underline{\underline{A}}\underline{d}_j = 0$ for $i \neq j$ and that $\underline{d}_i^T\underline{\underline{A}}\underline{d}_i = 1$. This shows that set $\{\underline{d}_1, \underline{d}_2, \dots, \underline{d}_n\}$ is $\underline{\underline{A}}$-conjugate by definition and that $\{\underline{d}_1, \underline{d}_2, \dots, \underline{d}_n\}$ is normal with respect to the $\underline{\underline{A}}$-norm. □

   (b) If $\underline{\underline{Q}} \in \mathbb{R}^{n\times n}$ is an orthogonal matrix ($\underline{\underline{Q}}^T\underline{\underline{Q}} = \underline{\underline{I}}$), $\underline{d}_1, \underline{d}_2, \dots, \underline{d}_n$ are $\underline{\underline{A}}$-conjugate vectors, and $\underline{\underline{D}} = [\underline{d}_1, \underline{d}_2, \dots, \underline{d}_n]$, show that the columns of $\underline{\underline{D}}\underline{\underline{Q}}$ are also $\underline{\underline{A}}$-conjugate vectors.

   Proof. Consider the matrix product $\left(\underline{\underline{D}}\underline{\underline{Q}}\right)^T\underline{\underline{A}}\left(\underline{\underline{D}}\underline{\underline{Q}}\right)$.

   $$\left(\underline{\underline{D}}\underline{\underline{Q}}\right)^T\underline{\underline{A}}\left(\underline{\underline{D}}\underline{\underline{Q}}\right)^T = \underline{\underline{Q}}^T\underline{\underline{D}}^T\underline{\underline{A}}\underline{\underline{D}}\underline{\underline{Q}}$$

   Since the columns of $\underline{\underline{D}}$ are $\underline{\underline{A}}$-conjugate

   $$= \underline{\underline{Q}}^T I \underline{\underline{Q}}$$
   $$= \underline{\underline{Q}}^T\underline{\underline{Q}}$$

   Since $\underline{\underline{Q}}$ is orthogonal

   $$= I$$

   This shows by part (a) that the columns of $\underline{\underline{D}}\underline{\underline{Q}}$ are $\underline{\underline{A}}$- conjugate. □

2. Page 162: Problem 6.3
   Verify that (6.19) and (6.17) are inverses of each other.

3. Page 162: Problem 6.4
   Use the Sherman Morrison formula (A.27) to show that (6.24) is the inverse of (6.25).

   Proof. The SR1 method has an update method for the matrix B_k given as

   $$B_{k+1} = B_k + \frac{(y_k - B_k s_k)(y_k - B_k s_k)^T}{(y_k - B_k s_k)^T s_k}$$

   In order to find an update formula for the inverse, H_{k+1} the Sherman-Morrison formula can be used, which states that if

   $$A_{k+1} = A_k + ab^T$$

then

$$A_{k+1}^{-1} = A_k^{-1} - \frac{A_k^{-1} a b^T A_k^{-1}}{1 + b^T A_k^{-1} a}$$

In our update formula the vectors a and b have the following values.

$$a = \frac{y_k - B_k s_k}{(y_k - B_k s_k)^T s_k}$$

and

$$b = y_k - B_k s_k$$

This means that the update formula for the inverse H_{k+1} can be found as follows. Note that H_k and B_k are inverses and that B_k and H_k are symmetrical.

$$\begin{aligned}
H_{k+1} &= H_k - \frac{1}{(y_k - B_k s_k)^T s_k} \frac{H_k (y_k - B_k s_k)(y_k - B_k s_k)^T H_k}{1 + (y_k - B_k s_k)^T H_k (y_k - B_k s_k)/((y_k - B_k s_k)^T s_k)} \\
&= H_k - \frac{(H_k y_k - s_k)(y_k^T H_k - s_k^T)}{(y_k - B_k s_k)^T s_k + (y_k - B_k s_k)^T (H_k y_k - s_k)} \\
&= H_k - \frac{(s_k - H_k y_k)(s_k - H_k y_k)^T}{(y_k^T - s_k^T B_k) s_k + (y_k^T - s_k^T B_k)(H_k y_k - s_k)} \\
&= H_k - \frac{(s_k - H_k y_k)(s_k - H_k y_k)^T}{y_k^T s_k - s_k^T B_k s_k + y_k^T H_k y_k - y_k^T s_k - s_k^T y_k + s_k^T B_k s_k} \\
&= H_k - \frac{(s_k - H_k y_k)(s_k - H_k y_k)^T}{y_k^T H_k y_k - y_k^T s_k} \\
&= H_k - \frac{(s_k - H_k y_k)(s_k - H_k y_k)^T}{(y_k^T H_k - s_k^T) y_k} \\
&= H_k + \frac{(s_k - H_k y_k)(s_k - H_k y_k)^T}{(s_k^T - y_k^T H_k) y_k} \\
&= H_k + \frac{(s_k - H_k y_k)(s_k - H_k y_k)^T}{(s_k - H_k y_k)^T y_k}
\end{aligned}$$

This last update formula is identical to equation (6.25), so this shows that (6.24) and (6.25) are indeed inverses of each other. □

4. Page 162: Problem 6.6
The square root of a matrix A is a matrix $A^{1/2}$ such that $A^{1/2} A^{1/2} = A$. Show that symmetric positive definite matrix A has a square root, and that this square root is itself symmetric and positive definite.

Proof. Let A be a symmetric positive definite matrix. Since A is symmetric positive definite this implies that the eigenvalues of A are real and positive and that the eigenvectors of A can be chosen to form an orthonormal basis of $\mathbb{R}^n$. This implies that A can be factored as UDU^T where D is a diagonal matrix with the eigenvalues of A on the diagonal and where U is an orthogonal matrix of the eigenvectors. I will define the matrix $D^{1/2}$ as the diagonal matrix whose diagonal entries are the positive square roots of the diagonal of D. This is well-defined as the eigenvalues are strictly greater than zero. This can also be shown as

$$D = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

$$D^{1/2} = \begin{bmatrix} \sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_n} \end{bmatrix}$$

$\square$

5. Implement the BFGS Method (Algorithm 6.1 on Page 140, or lecture notes).
6. Apply the BFGS method to the following function:

$$f(x, y) = (x^2 + y - 11)^2 + (x + y^2 - 7)^2.$$

Use $\underline{\underline{H}}_0 = \underline{\underline{I}}$ and an exact line search for each step length. Do all of the following:

-
-
-
-