Suggestion Answer for Homework 4, CS421

October 29, 1999

1. Denote

$$q_k(s) = \nabla f_k(s)^T s + \frac{1}{2} s^T \nabla^2 f_k s.$$

Since $q_k(0) = 0$, $\min_{||s|| \le \Delta_k} q_k(s) \le 0$. Suppose the *ratio* is finite. When the *ratio* is negative, $\min_{||s|| \le \Delta_k} q_k(s) < 0$. Then we have $f(x_k + s_k) > f(x_k)$

2.
We only need to prove $s_k^T y_k > 0$. Because

$$\nabla f(x_k + \alpha_k d_k)^T d_k \ge \beta \nabla f(x_k)^T d_k$$

then

$$\begin{aligned} s_k^T y_k = y_k^T s_k &= (\nabla f(x_k + \alpha_k d_k)^T - \nabla f(x_k)^T) s_k \\ &= (\nabla f(x_k + \alpha_k d_k)^T - \nabla f(x_k)^T) \alpha_k d_k \\ &\ge \alpha_k (\beta - 1) \nabla f(x_k)^T d_k \end{aligned}$$

Because $\beta < 1$, $\alpha > 0$ and $\nabla f(x_k)^T d_k < 0$, we have

$$s_k^T y_k > 0$$

3(a)

$$q(x_1, x_2) = x_1^2 + x_2$$

(b)

$$q(x_1, x_2) = x_1^2$$

When $X \in \{(0, x_2) : x_2 \in R^1\}$, $q(X)$ gets its global minimum. Since there are infinite points in this sets, there are infinitely many global minimizers.

4.

$$B_+ = M_+ M_+^T = (L+E)(L+E^T) = LL^T + EL^T + LE^T + EE^T$$

Because

$$LL^T = B$$

$$EL^T = \frac{(y-Lw)w^T L^T}{w^T w} = \frac{yw^T L^T - Lww^T L^T}{w^T w}$$

$$LE^T = \frac{Lw(y^T - w^T L^T)}{w^T w} = \frac{Lwy^T - Lww^T L^T}{w^T w}$$

$$EE^T = \frac{y-Lw}{w^T w} w^T w \frac{y^T - w^T L^T}{w^T w} = \frac{yy^T - yw^T L^T - Lwy^T + Lww^T L^T}{w^T w}$$

$$w^T w = \alpha^2 s^T L L^T S = \alpha^2 s^T B s$$

$$ww^T = \alpha^2 L^T s s^T L$$

$$\alpha^2 s^T B s = \frac{y^T s}{s^T B s} s^T B s = y^T s = s^T y$$

then

$$B_+ = B + \frac{yy^T - Lww^T L^T}{w^T w} = B + \frac{yy^T}{\alpha^2 s^T B s} - \frac{Bss^T B}{s^T B s} = B + \frac{yy^T}{s^T y} - \frac{Bss^T B}{s^T B s}$$

5.

```
x = -1:.05:1;
y = x;
z = zeros(41,41);
D = diag([2 -10]);
Q = orth(rand(2));
H = Q'*D*Q;
for i = 1:41
  for j = 1:41
    z(i,j) = .5 * [x(i),y(j)] * H * [x(i);y(j)];
  end
end
contour(x,y,z);
title('diag([2 -10])')
```

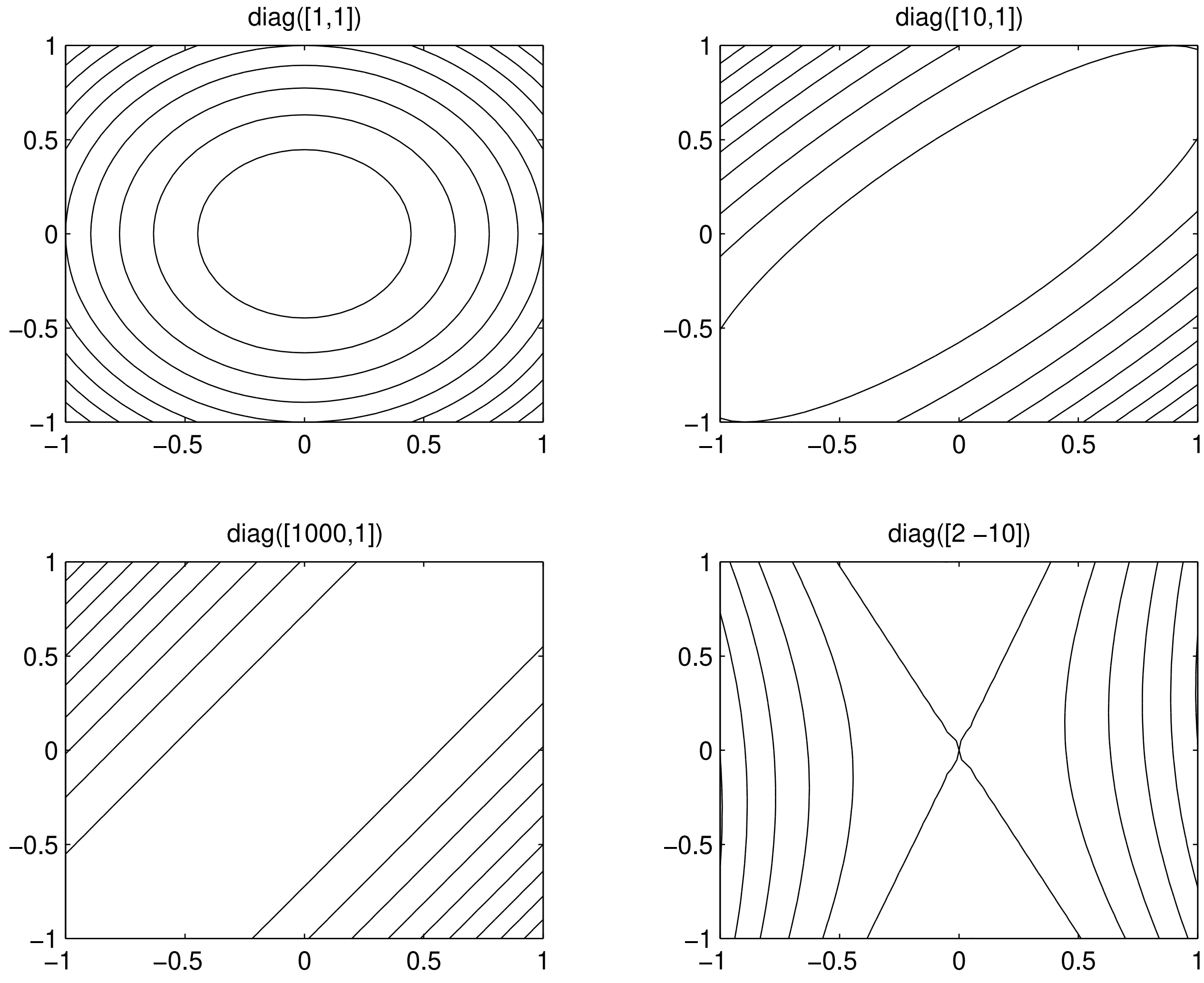

Figure 1: Contours

The first three contours are classes of ellipses because all eigenvalues are positive. When one eigenvalue becomes bigger and bigger, the focal length becomes bigger and bigger. Thus the contour near the origin looks like stright lines. The last one is a class of hyperbolas. This is because one eigenvalue is negative.

6.
Suppose

$$Q^T[s,r] = \begin{bmatrix} \alpha & \beta_1 \\ 0 & \beta_2 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix}$$

where α, β_1, β_2 are constants, Q is orthogonal matrix. We can assume $\alpha \neq 0$ because $\|s\| \neq 0$. Because $Cs = r$, then

$$(Q^T C Q)(Q^T s) = Q^T r$$

Let us denote $\tilde{C} = Q^T C Q$. Then

$$\tilde{C} \begin{bmatrix} \alpha \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Because C is symmetric, $\tilde{C}$ is also symmetric. Then

$$\tilde{C} = \begin{bmatrix} \beta_1/\alpha & \beta_2/\alpha & 0 & \cdots & 0 \\ \beta_2/\alpha & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ \vdots & * & * & \ddots & * \\ 0 & * & * & \cdots & * \end{bmatrix}$$

where '*'s can be any number which preserve symmetry.
Because orthogonal transformation will not change the Fubini norm, then

$$\|C\|_F = \|\tilde{C}\|_F$$

Thus the minimum Fubini norm can be achieved if only if all '*'s are zeros. Thus

$$C = Q \begin{bmatrix} \beta_1/\alpha & \beta_2/\alpha & 0 & \cdots & 0 \\ \beta_2/\alpha & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} Q^T$$

Then

$$C = \frac{Q\begin{bmatrix} 2\beta_1\alpha & \beta_2\alpha & 0 & \cdots & 0 \\ \beta_2\alpha & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}Q^T}{\alpha^2} - \frac{Q\begin{bmatrix} \beta_1\alpha^3 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}Q^T}{\alpha^4}$$

$\Longrightarrow$

$$C = \frac{rs^T + sr^T}{s^Ts} - (s^Tr)\frac{ss^T}{(s^Ts)^2}$$