APM462H1S, Winter 2014 ,   Assignment 3,

due: Monday March 17, at the beginning of the lecture.

Exercise 1. Solve problem 4 on page 282 of the textbook by following these steps. (For the notation, consult the book).

a. Show that there exists a vector d_1 which is Q-conjugate to d, and such that x_1 belongs to the plane spanned by d and d_1.

solution. If $x = cd$ for some c, we choose d_1 to be any nonzero vector that is Q-conjugate to d.

If x is not a multiple of d, we proceed would like to find a vector d_1 such that

$$x_1 = cd + d_1 \tag{1}$$

and

$$d^TQd_1 = 0. \tag{2}$$

To do this, multiply (1) on the left by d^TQ and rearrange to find that

$$c = -\frac{d_TQx_1}{d^TQd}.$$

So we define

$$d_1 = x_1 - \frac{d^TQx_1}{d^TQd}d.$$

Then it is straightforward to check that d_1 satisfies both (1) and (2).

(Note that this is basically a simple instance of the *Gram-Schmidt* procedure that you have seen in linear algebra.)

Note that the same argument shows that there exists a vector d_2 which is Q-conjugate to d, and such that x_2 belongs to the plane spanned by d and d_2.

b. Explain why x_1 can be found by starting at the origin, and taking two steps of the conjugate directions method, using directions d and d_1.

The "Expanding Subspaces Theorem" from the textbook and lecture states that if we start at the origin, and taking two steps of the conjugate directions method, using directions d and d_1, then the point that we reach will exactly be the minimum of f in the plane spanned by d and d_1, and this point is x_1.

Note that he same argument shows that x_2 can be found by starting at the origin, and taking two steps of the conjugate directions method, using the directions d and d_2.

c. Prove that $x_1 - x_2$ is Q-orthogonal to d.

If we find x_1 by taking two steps of the conjugate directions method, starting with the firection d, then the first step will take us to a point say $p = cd$ for some c (in fact $c = -\frac{d_TQx_1}{d^TQd}$ as above) and the second step will take be by a multiple of d_1 and will take us to x_1. So

$$x_1 = p + ad_1$$

for some a. (and in fact, $a = 1$.)

Similarly, to get to x_2, our first step will take us to *exactly the same point* p, since it is defined in exactly the same way as before. So

$$x_2 = p + bd_2$$

for some b. Thus $x_1 - x_2 = ad_1 - bd_2$, so

$$d^T Q(x_1 - x_2) = ad^T Qd_1 - bd^T Qd_2 = 0.$$

Exercise 2.
Minimize

$$f(x, y) = xy$$

subject to the constraints

$$x^2 + y^2 \leq 25, \qquad x + y \geq 1.$$

solution
let's write

$$g_1 = x^2 + y^2 - 25, \qquad g_2 = 1 - x - y,$$

so the constraints become

$$g_1(x, y) \leq 0, \qquad g_2(x, y) \leq 0.$$

We consider a number of cases.
case 1: no constraints active. Then the neessary conditions are $\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0$. This happens only at $(x, y) = (0, 0)$, which does not satisfy the constraints.

case 2: only g_1 active. Then the neessary conditions are are

$$\nabla f(x, y) + \mu \nabla g_1 = 0, \qquad \mu \geq 0$$

which is the same as the system of equations

$$y + 2\mu x = 0, \qquad x + 2\mu y = 0, \qquad g_1(x, y) = 0, g_2(x, y) < 0,$$

It is straightforward to solve and find that the only solution is $(x, y) = (5/\sqrt{2}, 5/\sqrt{2})$. Then $\mu = -1/2$, so this does not satisfy the necessary conditions.

case 3: only g_2 active. Then the neessary conditions are are

$$\nabla f(x, y) + \mu \nabla g_2 = 0, \qquad \mu \geq 0$$

which is the same as the system of equations

$$y - \mu = 0, \qquad x - \mu = 0, \qquad g_1(x, y) > 0, g_2(x, y) = 0,$$

It is straightforwrd to solve and find that the only solution is $(x, y) = (\frac{1}{2}, \frac{1}{2})$. Since $\mu = \frac{1}{2}$, this point satisfies the necessary conditions and is a candidate for the global minimizer. Here $f(x, y) = 1/4$.

case 4: g_1 and g_2 active. The only two solutions are $(x, y) = (4, -3)$ and $(x, y) = (-3, 4)$. At these points, $f(x, y) = -12$, which is smaller than the value of f at the other point found above. So the global minimum occurs at these two points and nowhere else.

remark. If we wanted to, we could have summarized the necessary conditions in the statement

$$\nabla f(x,y) + \mu_1 \nabla g_1 + \mu_2 \nabla g_2 = 0, \qquad \mu_1 g_1 + \mu_2 g_2 = 0, \qquad \mu_1, \mu_2 \text{ both positive.}$$

This is the Karush-Kuhn-Tucker optimality condition.

Exercise 3. Problem 10 on page 356 of the textbook.

solution We are given the relation

$$x(k+1) = \alpha x(k) - u(k), \qquad\qquad x(0) = F.$$

Then it is easy to check by induction that

$$x(k+1) = \alpha^{k+1} x(0) - \sum_{j=0}^{k-j} \alpha^j u(j)$$

for every k.

I will write $\vec{u} = (u(1), \dots, u(N))$, and from now on I will also write $(u_1, \dots, u_N)$ instead of $(u(1), \dots, u(N))$. So our goal is to maximize the function

$$J(\vec{u}) = \sum_{k=0}^{N} \psi(u_j) \beta^j$$

subject to the constraint $x(N+1) = 0$, which in view of the above can be written

$$h(\vec{u}) = \alpha^{N+1} x(0) - \sum_{j=0}^{N} \alpha^{N-j} u_j = 0.$$

a. The general optimality condition of for this problem is:

$$\nabla J(\vec{u}) + \lambda \nabla h(\vec{u}) = 0, \qquad\qquad h(\vec{u}) = 0.$$

(Here ∇ denotes the gradient with respect to variables $u_1, \dots, u_N$.) This can be written more explicitly as:

$$\psi'(u_j)\beta^j - \lambda \alpha^{N-j} = 0, \qquad j = 0, \dots, N.$$

b. In the special case $\psi(u) = u^{1/2}$, this becomes

$$\frac{1}{2}(u_j)^{-1/2}\beta^j + \lambda \alpha^{N-j} = 0, \qquad j = 0, \dots, N.$$

together with the constraint $h(\vec{u}) = 0$. We solve to find

$$u_j = (-2\lambda \alpha^{N-j} \beta^{-j})^{-2} \tag{3}$$

Substituting into the constraint, we find that

$$\alpha^{N+1} x(0) = \sum_{j=0}^{N} \frac{1}{4\lambda^2} \frac{\beta^{2j}}{\alpha^{N-j}}.$$

This can be solved to find that

$$\lambda = -\left(\frac{1}{\alpha^{N+1} x(0)} \sum_{j=0}^{N} \frac{1}{4} \frac{\beta^{2j}}{\alpha^{N-j}} \right)^{1/2} \tag{4}$$

(There are two choices for λ, and we can see from some of our earlier equations that the negative choice is the one we need.) Together, equations (3) and (4) specify the solution.

Exercise 4. Problem 12 on page 356 of the textbook.

Consider the quadratic program

$$\text{minimize} \quad f(x) = \frac{1}{2}x^T Qx - b^T x \qquad \text{subject to} \;\; Ax = c.$$

The point is, the set of points satisfying the constraint is a plane, and the restriction of f to this set is again a quadratic function. Hence its second derivatives are independent of the position in the plane. In particular, if it has a local minimum, then the matrix of second derivatives (in directions tangent to the contraint plane) is positive definite, so the (restricted) function is convex, and every critical point must be a global minimum.

Here is the same argument with more detail.

Let P denote the set of points satifsying the constraint, and let p be a point in P, and N be the nullspace of A. So $Ap = c$, and $x \in N$ if and only if $Ax = 0$. Thus

$$x \in N \quad \Longleftrightarrow \quad N(x+p) = c \quad \Longleftrightarrow \quad x + p \in P.$$

Now let $v_1, \ldots, v_k$ be a basis for N (where by definition k is the dimension of N.) Then every point in P can be written in a unique way in the form

$$p + y_1 v_1 + \ldots + y_k v_k$$

for some $y_1, \ldots, y_k$. If we write

$$V = \; n \times k \text{ matrix with columns } v_1, \ldots, v_k$$

and $y = (y_1, \ldots, y_k)$ (column vector) then we can rewrite this as

$$p + Vy$$

So studying minima etc of $f(x)$ subject to thte constraint $x \in P$ is the same as studying minima etc of the function

$$g(y) = f(p + Vy)$$

for $y \in E^k$. And by algebra,

$$\begin{aligned} g(y) &= \frac{1}{2}(p + Vy)^T Q(p + Vy) - b^T(p + Vy) \\ &= \frac{1}{2}y^T Ry - d^T y + e \end{aligned}$$

for

$$R = V^T QV, \qquad d = V^T b + p^T QV, \qquad c = \frac{1}{2}p^T Qp.$$

These are probably right, and it diesn't matter, because the point is that g is a quadratic function of k variables, and we don't need to know the exact formula for the coefficients.

Anyway, since g is quadratic, if it has a local minimum, then R is positive semidefinite, and this implies that g is convex, and hence that the local minimum is a global minimum.