APM 462: Nonlinear Optimization

Midterm Test
June 28, 2016.
Time: 90 minutes.

1. [10 points] Consider the problem

$$\begin{aligned} &\text{minimize:} && f(x,y,x) = x^{50} + \exp(y+x) + x^3y^{99} + \sin(z) \\ &\text{subject to:} && g(x,y,z) = x^2 - xy + y^2 + z^2 \leq 10, \end{aligned}$$

where f, g are fuctions from $\mathbb{R}^3$ to $\mathbb{R}$. Does this problem have a solution? Justify your answer. Hint: you may use the inequality $|ab| \leq \frac{1}{p}|a|^p + \frac{1}{q}|b|^q$.

Solution. The feasible set $\{(x,y,x) \in \mathbb{R}^3 \mid g(x,y,z) \leq 10\}$ is closed since g is continuous. We show the feasible set is also bounded, hence compact. Since a continuous function on a compact set attains its minimum, and f is continuous, the problem has a solution.

By the inequlity, $|xy| \leq \frac{1}{2}|x|^2 + \frac{1}{2}|y|^2$, since

$$\begin{aligned} 10 &\geq x^2 - xy + y^2 + z^2 \geq x^2 - |xy| + y^2 + z^2 \\ &\geq x^2 - (\frac{1}{2}|x|^2 + \frac{1}{2}|y|^2) + y^2 + z^2 \\ &\geq \frac{1}{2}(x^2 + y^2) + z^2, \end{aligned}$$

any (x,y,z) in the feasible set satisfies $\frac{1}{2}(x^2+y^2) + z^2 \leq 10$. So the feasible set is contained in the set $\{(x,y,x) \in \mathbb{R}^3 \mid \frac{1}{2}(x^2+y^2) + z^2 \leq 10\}$, which is bounded (being an ellipsoid).

2. [10 points] Consider the problem

$$\begin{aligned} &\text{minimize:} && f(x,y,x) = x^6y^5z^4 \\ &\text{subject to:} && h_1(x,y,z) = x^2 + y^2 + z^2 - 1 = 0 \\ & && h_2(x,y,z) = y^2 - z^2 = 0. \end{aligned}$$

a) What does it mean for a point (a,b,c) to be a *regular point* for the constraints?

b) Is the point $(1,0,0)$ a regular point for the constraints? Justify your answer.

Solution. **a**) A point (a,b,c) is a regular point for the constraints if $\nabla h_1(a,b,c) = (2a,2b,2c)$ and $\nabla h_2(a,b,c) = (0,2b,-2c)$ are linearly independent vectors.

b) Since $\nabla h_1(1,0,0) = (2,0,0)$ and $\nabla h_2(1,0,0) = (0,0,0)$ are linearly dependent, the point $(1,0,0)$ is not a regular point of the constraints.

3. [15 points] Assume that f is a convex function on $\mathbb{R}^n$, and that $L : \mathbb{R}^m \to \mathbb{R}^n$ is affine, which means that it has the form

$$L(x) = Bx + c$$

where B is an $n \times m$ matrix and $c \in \mathbb{R}^n$.

Let $g(x) = f(L(x))$, and prove that g is a convex function on $\mathbb{R}^m$.

Solution. Let $\alpha, \beta \geq 0$, and $\alpha + \beta = 1$. Note that $L(\alpha x + \beta y) = B(\alpha x + \beta y) + c = \alpha Bx + \beta By + c = \alpha(Bx + c) + \beta(By + c) = \alpha L(x) + \beta L(y)$.

By convexity of f we have, $g(\alpha x + \beta y) = f(L(\alpha x + \beta y)) = f(\alpha L(x) + \beta L(y)) \leq \alpha f(L(x)) + \beta f(L(y)) = \alpha g(x) + \beta g(y)$.

4 [15 points] Let f be a C^1 function on $\mathbb{R}^n$. Recall the method of steepest descent for minimizing f,

$$x_{k+1} := x_k - \alpha_k \nabla f(x_k).$$

Prove that $\nabla f(x_{k+1}) \cdot \nabla f(x_k) = 0$.

Solution. $f(x_{k+1}) = f(x_k - \alpha_k \nabla f(x_k)) = \min_{s \geq 0} f(x_k - s\nabla f(x_k))$, hence $0 = \frac{d}{ds}|_{s=\alpha_k} f(x_k - s\nabla f(x_k)) = \nabla f(x_{k+1}) \cdot (-\nabla f(x_k))$.

5. [15 points] Suppose that Q is a positive definite, symmetric, $n \times n$ matrix, and that $v \neq 0$ and b are vectors in $\mathbb{R}^n$. Solve the problem

$$\begin{aligned} \text{minimize:} \quad & f(x) = \frac{1}{2} x \cdot Qx - b \cdot x \\ \text{subject to:} \quad & h(x) = v \cdot x = 0. \end{aligned}$$

Here x denotes a point in $\mathbb{R}^n$. Your answer should express the minimum point x^* and a Lagrange multiplier λ in terms of Q, b and v. It might have the form: "$x^* = \ldots$, where λ satisfies $\ldots$,"

Solution. By the first order neccessary conditions for local minimum, x^* must satisfy: $\nabla f(x^*) + \lambda \nabla h(x^*) = (Qx^* - b) + \lambda v = 0$. Hence $x^* = Q^{-1}(b - \lambda v)$. To find λ, multiply both sides of the last equation by v: $0 = v \cdot x^* = v \cdot Q^{-1} b - \lambda v \cdot Q^{-1} v$. Solving for λ we get, $\lambda = \frac{v \cdot Q^{-1} b}{v \cdot Q^{-1} v}$. Checking 2nd order conditions at x^*, we have $\nabla^2 f(x^*) + \lambda \nabla^2 h(x^*) = Q + \lambda 0$

is positive definite, hence x^* is a strict local minimum. Since f is strictly convex, x^* is the unique global minimum.

6. [20 points] Consider the problem

$$\begin{aligned} \text{minimize:} \quad & f(x,y) = x^3 + y^2 \\ \text{subject to:} \quad & g(x,y) = (x+1)^2 + y^2 - 1 \leq 0. \end{aligned}$$

a) Show that all the points which satisfy the constraints are regular.

b) Find all points which satisfy the 1st order conditions for local minimum.

c) Which of the points you found in part b) satisfy the 2nd order necessary conditions? Justify your answer.

d) Which of the points you found in part b) is a local minimum? Justify your answer.

Solution. **(a)** If the constraint is inactive at (x,y) there is nothing to check, so (x,y) is regular. If (x,y) is active for the constraint, i.e. $g(x,y) = 0$, then since $\nabla g(x,y) = (2(x+1), 2y) \neq (0,0)$, the point (x,y) is regular.
(b) The 1st order conditions for local minimum give: $\nabla f(x) + \mu \nabla g(x) = (3x^2, 2y) + \mu(2(x+1), 2y) = 0$, where $\mu \geq 0$. This implies $2y(1+\mu) = 0$, hence either $y = 0$ or $\mu = -1$. Since μ is non-negative, it cannot equal -1. Since $g(x,y) = 0$, $y = 0$ implies that $x = 0$ or $x = -2$. So the candidates for local minimum point are $(0,0)$ and $(-2,0)$.
(c) The 1st order conditions above imply that $\mu = 0$ at $(0,0)$ and that $\mu = 6$ at $(-2,0)$. Note that $\nabla^2 f(x,y) = \begin{pmatrix} 6x & 0 \\ 0 & 2 \end{pmatrix}$ and $\nabla^2 g(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$. The tangent space to the constraint at $(0,0)$ is the line $x = 0$, and the tangent space to the constraint at $(-2,0)$ is the line $x = 0$ well.

So at $(0,0)$ we have, $\nabla^2 f(0,0) + 0\nabla^2 g(0,0) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$, and at $(-2,0)$ we have, $\nabla^2 f(-2,0) + 6\nabla^2 g(-2,0) = \begin{pmatrix} 0 & 0 \\ 0 & 14 \end{pmatrix}$. Both are positive semidefinite, so the 2nd order necessary conditions are satisfied (we don't need to restrict attention to the tanget space in this case since the matrices are positive semidefinite).
(d) The point $(0,0)$ is not a local minimum: if we move a bit in the direction $(0,-1)$ the objective function f is getting smaller. The point $(-2,0)$ is a regular point and the constraint $g(x,y) = 0$ is strongly active at this point since $\mu = 6 > 0$. The tanget space to the strongly active constraint at $(-2,0)$ is the line $x = 0$ so the 2nd order sufficient condition is $(0,y)[\nabla^2 f(-2,0) + 6\nabla^2 g(-2,0)] \begin{pmatrix} 0 \\ y \end{pmatrix} = (0,y) \begin{pmatrix} 0 & 0 \\ 0 & 14 \end{pmatrix} \begin{pmatrix} 0 \\ y \end{pmatrix} = 14y^2 > 0$ for all $(0,y) \neq (0,0)$, hence $(-2,0)$ is a strict local minimum.

7. [15 points] Fix a point $y \in \mathbb{R}^n$ and let $f : \mathbb{R}^n \to \mathbb{R}$ be the square of the distance to y, that is $f(x) := ||y - x||^2 = (y_1 - x_1)^2 + \cdots + (y_n - x_n)^2$. Let $g : \mathbb{R}^n \to \mathbb{R}$ be a convex C^1 function.

a) Prove that the function f is strictly convex.

b) Suppose x^* solves the problem

$$\begin{array}{ll} \text{minimize:} & f(x) \\ \text{subject to:} & g(x) \leq 0. \end{array}$$

Prove that x^* is unique.

c) Prove that $y - x^*$ is parallel to $\nabla g(x^*)$, where x^* is the solution from part b).

Solution. **a**) $\nabla^2 f(x) = 2I > 0$. Here I is the identity $n \times n$ matrix.
b) Note that since g is convex, its zero sublevel set $Z := \{x \mid g(x) \leq 0\}$ is a convex set. Let $f(x^*) = d$. Suppose x^* is not unique, that is suppose there exists two distinct points x_1 and x_2 in Z such that $f(x_1) = f(x_2) = d$. Since Z is convex, $\alpha_1 x_1 + \alpha_2 x_2 \in Z$ for all non-negative α_1 and α_2 satisfying $\alpha_1 + \alpha_2 = 1$. Since f is strictly convex, $d \leq f(\alpha_1 x_1 + \alpha_2 x_2) < \alpha_1 f(x_1) + \alpha_2 f(x_2) = \alpha_1 d + \alpha_2 d = d$, a contradiction.
c) By 1st order neccessary condition $0 = \nabla f(x^*) + \mu \nabla g(x^*) = 2(x^* - y) + \mu \nabla g(x^*)$. Hence $y - x^* = \frac{\mu}{2} \nabla g(x^*)$.