Assignment 2

APM462 – Nonlinear Optimization – Summer 2016

Christopher J. Adkins

Solutions

Question 1 To approximate a function $g : [0,1] \to \mathbb{R}$ by an n-th order polynomial, one can minimize the function f defined by

$$f(a) = \int_0^1 (g(x) - p_a(x))^2 dx$$

where, for $a = (a_0, \ldots, a_n) \in \mathbb{R}^{n+1}$, we use the notation

$$p_a(x) = a_0 + a_1 x + \ldots + a_n x^n = (\mathbf{x}, a) \quad \& \quad \mathbf{x} = (1, x, \ldots, x^n)$$

In HW1 you investigated the special case when $g(x) = x^2$ and $n = 1$. Here we continue to investigate this problem in more generality.

(a) Show that $f(a)$ can be written in the form

$$f(a) = a^T Q a - 2b^T a + c$$

for a $(n+1) \times (n+1)$ matrix Q, a vector $b \in \mathbb{R}^{n+1}$ and a number c. Find formulas for Q, b and c. It should be clear from your formula that Q is symmetric.

(b) Find the first-order necessary condition for a point $a_* \in \mathbb{R}^2$ to be a minimum point for f. $\square$

(c) For $n = 0, 1, 2$ show that the matrix Q in part b) is invertible. Assuming that Q is invertible for any n, conclude that f can have at most one local minimum.

(d) For $n = 0, 1, 2$ is Q positive semidefinite? positive definite? Is a_* a local minimum?

Solution

(a) Expanding out $f(a)$, we see that

$$\begin{aligned} f(a) = \int_0^1 (g(x) - p_a(x))^2 dx &= \int_0^1 (g(x) - (\mathbf{x}, a))^2 dx \\ &= \int_0^1 \left[(\mathbf{x}, a)(\mathbf{x}, a) - 2g(x)(\mathbf{x}, a) + g(x)^2\right] dx \\ &= \int_0^1 \left[(a, \mathbf{x}\mathbf{x}^T a) - 2(g(x)\mathbf{x}, a) + g(x)^2\right] dx \\ &= (a, Qa) - 2(b, a) + c \end{aligned}$$

where

$$Q = \int_0^1 \mathbf{x}\mathbf{x}^T dx \implies Q_{ij} = \int_0^1 x^i x^j dx = \frac{1}{i+j+1}$$

$$b = \int_0^1 g(x)\mathbf{x} dx \implies b_i = \int_0^1 g(x) x^i dx$$

$$c = \int_0^1 g(x)^2 dx$$

(b) As usual with functions of this form, we know that the first order condition is given by $Qa_* = b$

(c) The particular form of Q implies that it is invertible since each column is linearly independent from one another. Let's do a quick check to try and find a constants $\alpha_1, \alpha_2 \in \mathbb{R} \setminus \{0\}$ such that $\alpha_1 Q_j + \alpha_2 Q_{j'} = 0$ where Q_j is the j-th column of Q.

$$\alpha_1 Q_j + \alpha_2 Q_{j'} = 0 \implies \frac{\alpha_1}{i+j+1} + \frac{\alpha_2}{i+j'+1} = 0 \implies \alpha_1 j' + \alpha_2 j + i(\alpha_1 + \alpha_2) + \alpha_1 + \alpha_2 = 0$$

We need to equation to work for all $i \leqslant n$, but that means $\alpha_1 = -\alpha_2$ so the relation reduces to

$$\alpha_1 (j - j') = 0$$

but $j \neq j'$ which shows that $\alpha_1 = \alpha_2 = 0$. Thus each column is linearly independent.

(d) The particular form of Q implies that it is positive definite since we have that

$$(a, Qa) = \int_0^1 p_a(x) dx \geqslant 0 \quad \& \quad \det Q \neq 0 \quad \text{(invertible)}$$

Thus a_* is a local minimum. □

Question 2 For $x, y \in \mathbb{R}$, define $f(x,y) = xy$ and $h(x,y) = x^2 + y^2 - 10$. Consider the optimization problem:

$$\text{minimize } f(x,y) \quad \text{subject to } h(x,y) = 0$$

(a) Draw a picture and guess the local minimum and maximum points for this problem.

(b) Show that every feasible point is regular. (Recall that a point is called feasible if it satisfies the constraints.)

(c) Use the first order necessary condition to find all candidates for local minimum points. (You should get four candidates)

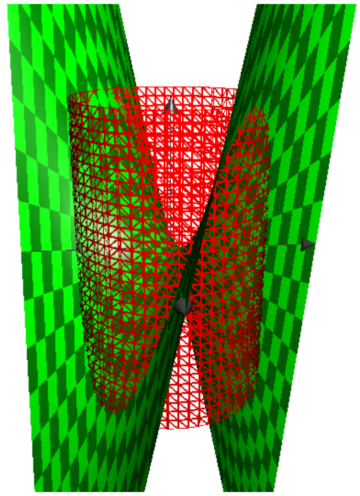
(d) Compute the tangent spaces to all the candidates on the circle

$$M := \{(x,y) \in \mathbb{R}^2 : h(x,y) = 0\}$$

(e) Use the second order condition to determine which of the candidates are local minimum points.

Solution

(a) The picture should look something:



it's easy to see the max and min's are located at $(\pm\sqrt{5}, \pm\sqrt{5})$

(b) All the points are regular since

$$\nabla h(x,y)\Big|_{h=0} = (2x, 2y)_{h=0} \neq 0$$

(c) The first order condition is given by

$$\nabla f = (y,x) = 2\lambda(x,y) = \lambda \nabla h \implies \begin{cases} y = 2\lambda x \\ x = 2\lambda y \end{cases}$$

solving the system gives

$$x = 4\lambda^2 x \implies \lambda = \pm\frac{1}{2} \implies x = \pm y \quad \& \quad y = \pm x$$

Then using the constraint shows

$$h(x,y) = x^2 + x^2 - 10 = 0 \implies x^2 = y^2 = 5$$

$$\mathbf{x}_* = \underbrace{(\sqrt{5}, \sqrt{5}), (-\sqrt{5}, -\sqrt{5})}_{\lambda=1/2} \quad \& \quad \underbrace{(\sqrt{5}, -\sqrt{5}), (-\sqrt{5}, \sqrt{5})}_{\lambda=-1/2}$$

are the four candidates.

(d) The tangent space is given by solutions to $(\nabla h(\mathbf{x}_*), v) = 0$. Thus

$$T_{\mathbf{x}_*}M = \{v \in \mathbb{R}^2 : (\nabla h(\mathbf{x}_*), v) = 0\} = \begin{cases} \text{span}(1,1) & \lambda = -1/2 \\ \text{span}(-1,1) & \lambda = 1/2 \end{cases}$$

(e) The second order condition is given by $(v, \nabla^2(f - \lambda h)v)_{x=\mathbf{x}_*} \geqslant 0$ when $v \in T_{\mathbf{x}_*}M$. In our case we see

$$\nabla^2(f - \lambda h) = \begin{pmatrix} -2\lambda & 1 \\ 1 & -2\lambda \end{pmatrix}$$

It's easy to see when $\lambda = -1/2$, i.e. $v \in T_{\mathbf{x}_*}M$

$$(v, \nabla^2(f - \lambda h)v) \geqslant 0$$

and when $\lambda = 1/2$ we have $(v, \nabla^2(f - \lambda h)v) \leqslant 0$ Thus. $\lambda = -1/2$ produces local minima. □

Question 3 Let

$$f(\mathbf{x}) = \frac{1}{2}(\mathbf{x}, Q\mathbf{x}) - (b, \mathbf{x})$$

where Q is positive semidefinite $n \times n$ symmetric matrix and $x \in \mathbb{R}^n$.

(a) Prove that $f(x)$ is a convex function.

(b) Prove that $g(x, y) = 4x^2 - 14x + 7y^2 + 8$, where $(x, y) \in \mathbb{R}^2$, is a convex function.

(c) Find the maximum of g on the set $\Omega := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leqslant 9\}$

Solution

(a) To check if f is convex it suffices for $\nabla^2 f$ to be positive semidefinite since f is C^2. We see

$$\nabla^2 f = Q$$

which is given by assumption of Q.

(b) Again, it suffices for $\nabla^2 g$ to be positive semidefinite since g is C^2. We see

$$\nabla^2 g = \begin{pmatrix} 8 & 0 \\ 0 & 14 \end{pmatrix}$$

which is clearly positive definite since the eigenvalues are positive.

(c) Since g is convex, we know the maximum is obtained on the boundary of $\Omega := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leqslant 9\}$. For the sake of a reminder, let's convert this back to a 1-d calculus problem using $x = 3\cos\theta$ and $y = 3\sin\theta$. We only need to maximize in terms of θ now. We see

$$f(x, y)\Big|_{\Omega} = 36\cos^2\theta - 42\cos\theta + 63\sin^2\theta + 8 = 27\sin^2\theta - 42\cos\theta + 44$$

The critical points are given by

$$f'(\theta) = 54\sin\theta\cos\theta + 42\sin\theta = 0 \implies \cos\theta = -\frac{7}{9} \quad \& \quad \sin\theta = 0$$

It's clear that $\sin\theta = 0$ will produce a minimum since $f''((n + 1/2)\pi) > 0$. Thus the maximum θ is given by $\cos\theta_* = -7/9$:

$$f(\theta_*) = 27\left(1 - \cos^2\theta_*\right) - 42\cos\theta_* + 44 = \frac{262}{3}$$

□

Question 4 A cardboard box is to be manufactured. The top, bottom, front, and back faces must be double weight (i.e., two pieces of cardboard). A problem posed is to find the dimensions of such a box that maximize the volume for a given amount of cardboard, equal to $A > 0$ square meters. Let us denote width, length, and height of the box by x, y, and z, respectively.

(a) Write the problem as a minimization problem and show that all feasible points of the constraints are regular.

(b) Find the point (x, y, z) which satisfies the 1st order necessary condition for a minimum

(c) Verify the 2nd order condition for a minimum.

Solution

(a) We want to maximize the volume V, given some area A. We know the classic formulas to calculate volume and area of a box:

$$V(x,y,z) = xyz \quad \& \quad A = 4xy + 2yz + 4zx \implies h(x,y,z) = 4xy + 2yz + 4zx - A$$

Thus the problem we want to solve is

$$\text{minimize } -V(x,y,z) \quad \text{subject to } h(x,y,z) = 0$$

When checking the gradient of the constraint, we see

$$\nabla h\Big|_{h=0} = 2\begin{pmatrix} 2y+2z \\ 2x+z \\ 2x+y \end{pmatrix}_{h=0} \neq 0$$

thus all feasible points are regular.

(b) In this case, we have the first order condition of

$$\nabla(-V - \lambda h) = 0 \implies \begin{cases} -yz - 2\lambda(2y+2z) = 0 \\ -xz - 2\lambda(2x+z) = 0 \\ -xy - 2\lambda(2x+y) = 0 \end{cases}$$

By adding all the equations together we see

$$A = -16\lambda(2x+y+z)$$

Now by symmetry, we must have $y = z$. The first condition now gives us

$$-z(z+8\lambda) = 0 \implies y = z = -8\lambda$$

and the second or third gives

$$x = -4\lambda$$

Now we see

$$A = 384\lambda^2 \implies \lambda = -\frac{1}{8}\sqrt{\frac{A}{6}}$$

note we threw away the positive root since the lengths are positive. Thus the only critical point is

$$\mathbf{x}_* = \sqrt{\frac{A}{6}}\left(\frac{1}{2}, 1, 1\right)$$

(c) The 2nd order condition is now easy to check. We compute:

$$\nabla^2(-V - \lambda h)_{\mathbf{x}=\mathbf{x}_*} = \begin{pmatrix} 0 & -z-4\lambda & -y-4\lambda \\ -z-4\lambda & 0 & -x-2\lambda \\ -y-4\lambda & -x-2\lambda & 0 \end{pmatrix}_{\mathbf{x}=\mathbf{x}_*} = 2\lambda\begin{pmatrix} 0 & 2 & 2 \\ 2 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix}$$

Note that

$$\nabla h(\mathbf{x}_*) = 2(-32\lambda, -16\lambda, -16\lambda) = -32\lambda(2,1,1)$$

which allows us to note the tangent space is given by

$$T_{\mathbf{x}_*}M = \{v \in \mathbb{R}^3 : (\nabla h(\mathbf{x}_*), v = 0)\} = \text{span}\left\{ \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -2 \\ 2 \\ 2 \end{pmatrix} \right\} \implies Z(\mathbf{x}_*) := \begin{pmatrix} 0 & -2 \\ 1 & 2 \\ -1 & 2 \end{pmatrix}$$

we now see the second order condition is given by

$$Z^T \nabla^2 (-V - \lambda h) Z = \begin{pmatrix} -4\lambda & 0 \\ 0 & -16\lambda \end{pmatrix}$$

which is positive definite since $\lambda < 0$. Thus $\mathbf{x}_*$ is a local minimum.