APM462H1S, Winter 2014, About the Final

solutions to a selected few sample problems

minimization with equality constraints

Thesample problems are all solved using Lagrange multipliers.

You can probably find solutions to some of them (the ones with actual numbers) using Wolfram Alpha.

minimization with inequality constraints (and maybe equality constraints as well)

In general, the procedure for these is to identify all possible critical points by considering different cases, when different combinations of constraints are active. Then you can try to find out which (if any) are global minimum points, for example by comparing the values of the function f that you are trying to minimize (although sometimes a little more thought is needed.)

second-order conditions for constrained minimization problems.

We did not do many of thses, so here are some detailed solutions:

- Consider the problem

$$\begin{aligned} &\text{minimize } f(x,y,z) := \frac{1}{2}(x^2+y^2-z^2) \\ &\text{subject to } h(x,y,z) := x+2y+3z-4=0. \end{aligned}$$

  Find a point satisfying the first-order necessary conditions, and determine whether or not it is a local minimum.

  solution: First, find a point satisfying the first-order necessary conditions. Using Lagrange multipliers, we have to solve the system

$$\begin{aligned} x+\lambda &= 0 \\ y+2\lambda &= 0 \\ -z+3\lambda &= 0 \\ h(x,y,z) &= 0\,. \end{aligned}$$

  We easily find that

$$\vec{x}^* = (x,y,z) = (-1,-2,3), \qquad \lambda = 1.$$

  For the second-order conditions, we form the matrix

$$L = \nabla^2 f(\vec{x}) + \lambda \nabla^2 h(\vec{x}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

  We have to check whether $\vec{v}^T\, L\, \vec{v} \geq 0$ (or maybe > 0) for all $\vec{v}$ in $M :=$ the tangent space to the surface defined by the constraint.

  In this case (and in most problems we will encounter),

$$\begin{aligned} M &= \{\vec{v} = (v_1,v_2,v_3) \in \mathbb{R}^3 : \nabla h(\vec{x}^*)(\vec{v}) = 0\} \\ &= \{\vec{v} = (v_1,v_2,v_3) \in \mathbb{R}^3 : v_1+2v_2+3v_3 = 0\}. \end{aligned}$$

  There are several ways to do this, but here is the standard procedure:

  step 1. choose a basis for M.

We just need to find two linearly independent vectors in M. We can do this by inspection, for example

$$\vec{v}_1 = (2, -1, 0), \qquad \vec{v}_2 = (3, 0, -1)$$

step 2. Write the 2×2 matrix whose i, j entry is $v_i^T M v_j$. This is

$$\begin{pmatrix} 6 & 6 \\ 6 & 8 \end{pmatrix}$$

step 3. determine whether it is positive (semi)-definiite.

Since the diagonal entries are positive and the determinant is positive, the matrix is positive definite. (This only works for 2×2 matrices.) Thus $\vec{x}$ is a local minimum.

remark: If

$$M = \{\vec{v} = (v_1, v_2, v_3) \in \mathbb{R}^3 : av_1 + bv_2 + cv_3 = 0\},$$

then one possible basis for M is always

$$\vec{v}_1 = (b, -a, 0), \qquad \vec{v}_2 = (c, 0, -a).$$

- Consider the problem

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

Find a point satisfying the first-order necessary conditions, and determine whether or not it is a local minimum.

solution: First, find a point satisfying the first-order necessary conditions. Using Lagrange multipliers, we have to solve the system

$$\begin{aligned} 1 + \lambda x &= 0 \\ 2 + \lambda y &= 0 \\ 3 - \lambda z &= 0 \\ h(x,y,z) &= 0\,. \end{aligned}$$

Note that $\lambda = 0$ will not work, so we can assume that λ is nonzero. Then we write $(x,y,z) = (1,2,-3)/\lambda$ and adjust λ until the constraint is satisfied. If we try to do this, we see that in fact the constraint can never be satisfied. So there is no point (x,y,z) satisfying the first-order conditions, and the problem cannot be solved. oops!

- Here is a revised version of the above problem, which has been changed to make it possible to find a solution.

Consider the problem

$$\begin{aligned} &\text{minimize } f(x,y,z) := x + 2y + 3z \\ &\text{subject to } h(x) := \frac{1}{2}(x^2 + y^2 - z^2) + 20 = 0. \end{aligned}$$

Find a point satisfying the first-order necessary conditions, and determine whether or not it is a local minimum.

The solution can be found on the next page. I strongly recommend that you try to solve it before looking at the solution.

solution: Using Lagrange multipliers, we have to solve the system

$$\begin{aligned} 1 + \lambda x &= 0 \\ 2 + \lambda y &= 0 \\ 3 - \lambda z &= 0 \\ h(x, y, z) &= 0\,. \end{aligned}$$

Note that $\lambda = 0$ will not work, so we can assume that λ is nonzero. Then we write $(x, y, z) = (1, 2, -3)/\lambda$ and adjust λ until the (new) constraint is satisfied. This leads to:

$$\vec{x}_1 = (2, 4, -6), \quad \lambda = -\frac{1}{2} \qquad \text{or} \qquad \vec{x}_2 = (-2, -4, 6), \quad \lambda = \frac{1}{2}.$$

(The problem only asks us to find and examine one point, but let's consider both.)

For the second-order conditions, we form the matrix

$$L = \nabla^2 f(\vec{x}_i) + \lambda \nabla^2 h(\vec{x}_i) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

at $i = 1, 2$.

At the point $\vec{x}_1$: Here the above matrix is:

$$L_1 = -\frac{1}{2}\nabla^2 h(\vec{x}_1) = -\frac{1}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

We have to check whether $\vec{v}^T\, L\, \vec{v} \geq 0$ (or maybe > 0) for all $\vec{v}$ in $M :=$ the tangent space to the surface defined by the constraint.

In this case (and in most problems we will encounter),

$$\begin{aligned} M &= \{\vec{v} = (v_1, v_2, v_3) \in \mathbb{R}^3 : \nabla h(\vec{x}^*)(\vec{v}) = 0\} \\ &= \{\vec{v} = (v_1, v_2, v_3) \in \mathbb{R}^3 : 2v_1 + 4v_2 + 6v_3 = 0\}. \end{aligned}$$

This is the same M as in the first problem, and L_1 differs from the earlier L only by the factor of $-1/2$.

It follows that, from what we did above, for any $\vec{y}$ in M,

$$\vec{y}^T L_1 \vec{y} = (-1/2)\ \vec{y}^T \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \vec{y} \leq 0.$$

(In fact it is < 0 unless $\vec{y} = 0$.)

Thus $\vec{x}_1$ is not a local minimum – on the contrary, it is a local maximum.

At the point $\vec{x}_2$: Here the above matrix is:

$$L_2 \;=\; \frac{1}{2}\nabla^2 h(\vec{x}_1) = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

So by similar reasoning, again using what we know from the earlier exercise, we can find that $\vec{x}_2$ is a local minimum.

Calculus of variations:

The questions from the earlier handout were rather calculus of variations problems, of a sort that we have seen before.

Systems of ODEs and matrix exponentials.

The questions from the earlier handout were mostly quite routine. The hardest one, I believe, was the last one:

- Let

$$M = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

a compute e^{tM}.

solution:

We compute

$$M^2 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad M^3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad M^4 = I,$$

Then $M^{j+4k} = M^j$ for every positive integer k and $j \in \{0, 1, 2, 3\}$.

From this one can check that

$$e^{tM} = \begin{pmatrix} \cos t & \sin t & 0 \\ -\sin t & \cos t & 0 \\ 0 & 0 & e^t \end{pmatrix}.$$

b. Write down a formula for the solution of

$$\frac{d}{dt}x(t) = Mx(t) + \begin{pmatrix} 0 \\ \alpha(t) \\ 0 \end{pmatrix}.$$

solution: Let's suppose that the initial condition is $x(0) = x^0$. (The problem should probably have specified this.) Then we know that the general solution formula is

$$x(t) = e^{tM}x^0 + e^{tM} \int_0^t e^{-sM} \begin{pmatrix} 0 \\ \alpha(s) \\ 0 \end{pmatrix} ds$$

where e^{tM} and e^{-sM} can be found in the solution to part **a**. I guess that there is not much point in writing the whole thing out.

controllability

- Consider the equation
$$\frac{d}{dt}x(t) = Mx(t) + N\alpha(t), \qquad \alpha(t) \in [-1,1] \text{ for all } t$$
where
$$M = \begin{pmatrix} 0 & 1 & 0 & 3 & 0 \\ 1 & 2 & 0 & 4 & 0 \\ 0 & 2 & -1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}, \qquad N = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$
Is this system controllable?

  solution

  For every $k = 0, 1, \ldots$, $M^k N$ has the form
$$N = \begin{pmatrix} * \\ * \\ * \\ * \\ 0 \end{pmatrix}.$$
So the controllability matrix G has the form
$$G = [N, MN, \ldots, N^4 M] = \begin{pmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$
Here * stands for some number that whose value we do not care about. Thus its rank is at most 4, so the system is not controllable.

- Same question, same M, but
$$N = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}.$$

  solution. This problem is too hard for a test. But the asnwer is no. One way to see this is to first to check that e^{tM} has the form
$$e^{tM} = \begin{pmatrix} * & * & * & * & 0 \\ * & * & * & * & 0 \\ * & * & * & * & 0 \\ * & * & * & * & 0 \\ 0 & 0 & 0 & 0 & e^{3t} \end{pmatrix},$$
Also, recall that x^0 is in the reachable set if there is some admissible control and some time t such that
$$x^0 = -\int_0^t e^{-sM} N\alpha(s)\, ds = -\int_0^t \begin{pmatrix} * \\ * \\ * \\ * \\ e^{-3s}\alpha(s) \end{pmatrix} ds,$$
using what we know about the form of e^{-sM} and N. Now the problem can be finished like one of the homework exercises.

Pontryagin's Maximum Principle

- Consider the equation

$$\frac{d}{dt}x_1(t) = x_2(t) + \alpha_1(t),$$
$$\frac{d}{dt}x_2(t) = -x_1(t) + \alpha_2(t),$$

with

$$\alpha_1^2(t) + \alpha_2^2(t) \le 1 \quad \text{for all } t. \tag{1}$$

(Note that this is different from the usual constraint, which would be $\alpha(t) \in [-1,1]^2$ for all t.)

Suppose $\alpha^*(\cdot)$ is an optimal control, steering the system from x^0 to the origin in the minimum possible time τ^*.

Note: This problem is too hard for an exam, which is one reason it is not on the Final.

a. What can we deduce about $\alpha^*(t)$ from the Pontryagin maximum principle?

solution. I will write $\alpha(t)$ below, instead of $\alpha^*(t)$, because it's easier.

(It is a fact that the Pontryagin maximum principle still holds with the constraint (1), although it has to be modified a little bit. There's only one reasonable choice for the correct modification, and you can probably figure it out.)

First let's rewrite the system in the standard way:

$$\frac{d}{dt}x(t) = Mx(t) + N\alpha(t), \qquad \alpha_1^2(t) + \alpha_2^2(t) \le 1 \text{ for all } t$$

where

$$M = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad N = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

It is easy to check that the system is controllable.

Pontryagin's maximum principle says that there exists a vector $h = \binom{h_1}{h_2}$ such that for every t from 0 to τ^*,

$$h^T e^{-tM} N\alpha(t) = \max_{\{a: a_1^2 + a_2^2 \le 1\}} h^T e^{-tM} Na. \tag{2}$$

Since N is the identity matrix, this is the same as

$$h^T e^{-tM}\alpha(t) = \max_{\{a: a_1^2 + a_2^2 \le 1\}} h^T e^{-tM} a.$$

Since $b^T a$ is just the dot product of b and a, and since $(h^T e^{-tM})^T = (e^{-tM})^T h = e^{tM} h$, this is the same as

$$(e^{tM}h)\cdot\alpha(t) = \max_{\{a: a_1^2 + a_2^2 \le 1\}} (e^{tM}h)\cdot a.$$

Clearly, since a can be any vector with length at most 1, the max on the right-hand side occurs when $|a| = 1$ and a is parallel to $e^{tM}h$, ie

$$\text{max occurs when } a = \frac{e^{tM}h}{|e^{tM}h|}, \quad \text{which in fact equals } e^{tM}\frac{h}{|h|}.$$

It follows that for every $t \in [0, \tau^*]$,

$$\alpha(t) = e^{tM} \tilde{h}, \qquad \text{for } \tilde{h} = \frac{h}{|h|}.$$

If we want to, we could write this out more explicitly, since we know that

$$e^{tM} = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix},$$

but let's leave it as it is.

b. What is the optimal control? Could you have figured this out without using the Pontryagin maximum principle?

Let's substitute what we know about $\alpha(t)$ into the formula for the solution of the ODE:

$$\begin{aligned} x(t) &= e^{tM} x^0 + \ e^{tM} \int_0^t e^{-sM} \alpha(s) \ ds \\ &= \ e^{tM} \left[x^0 + \int_0^t e^{-sM} e^{sM} \tilde{h} \right] \\ &= \ e^{tM} \left[x^0 + t\tilde{h} \right]. \end{aligned}$$

The way to make this equal zero as quickly as possible is to choose (indeed, the only way to ever make it equal zero) is to choose $\tilde{h} = -\frac{x^0}{|x^0|}$.

Then

$$x(t) = e^{tM} [(1 - \frac{t}{|x^0|}) x^0] \ = \ (1 - \frac{t}{|x^0|}) \ e^{tM} x^0$$

and

$$\alpha(t) = -\frac{1}{|x^0|} e^{tM} x^0 \ = \ -\frac{x(t)}{|x(t)|}.$$

So $\alpha(t)$ is always the unit vector pointing directly from $x(t)$ toward the origin. The is exactly what you would guess, if you just thought about the problem without doing any computations: if we want to get toward the origin as quickly as possible, just use the control to push $a(t)$ toward the origin.

- Consider the equation

$$\frac{d}{dt} x(t) = Mx(t) + N\alpha(t), \qquad \alpha(t) \in [-1, 1] \text{ for all } t$$

where

$$M = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \qquad N = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Suppose $\alpha^*(\cdot)$ is an optimal control, steering the system from x^0 to the origin in the minimum possible time. What can we deduce about $\alpha^*(t)$ from the Pontryagin maximum principle?

solution Without typing out all the details, the Pontryagin maximum principle says that there exists a vector $h = (h_1, h_2, h_3)^T$ such that

$$(\frac{1}{2}h_1t^2 + h_2t + h_3)\alpha(t) \;=\; \max_{|a|\leq 1}(\frac{1}{2}h_1t^2 + h_2t + h_3)a.$$

It follows that

$$\alpha(t) = \mathrm{sign}(\frac{1}{2}h_1t^2 + h_2t + h_3).$$

So $\alpha(t) = \pm 1$ fot every t that is not a root of the polynomial $\frac{1}{2}h_1t^2+h_2t+h_3$, and it can only change sign when y is a root (which happens for at most 2 values of t.)