Problem 1: Constrained optimization

In what follows, you may accept that $f(x, y) = xy$ attains maximal and minimal values on the curve $x^2 - xy + y^2 = 3$.

(a) Use the method of Lagrange multipliers to find these extreme values and the point(s) where they are attained.

(b) The quadratic formula allows you to solve for y in terms of x on the curve: $y(x) = (x \pm \sqrt{x^2 - 4(x^2 - 3)})/2 = (x \pm \sqrt{12 - 3x^2})/2$ (with $|x| \leq 2$ so that the square root makes sense). Hence, we could instead try to find the extreme values for $f(x, y(x)) = x \cdot y(x) = (x^2 \pm x\sqrt{12 - 3x^2})/2$ for $-2 \leq x \leq 2$ via single-variable calculus. Is that more or less appetizing than the method in (a)?

Solution:

(a) Let $g(x, y) = x^2 - xy + y^2$, so we seek the local extrema for f on the level set $g = 3$ and then will compare the values of f at those points to see which are largest and smallest.

The method of Lagrange multipliers says that if (a, b) is a point on the level curve where f attains a local extremum then $(\nabla f)(a, b)$ is a scalar multiple of $(\nabla g)(a, b)$ provided that $(\nabla g)(a, b) \neq \mathbf{0}$. So our first step is to compute the gradients of f and g at a general point (x, y) and to figure out where (if anywhere) ∇g vanishes on the level set $g = 3$ (such bad points will have to be analyzed separately).

We compute $\nabla f = \begin{bmatrix} y \\ x \end{bmatrix}$ and $\nabla g = \begin{bmatrix} 2x - y \\ -x + 2y \end{bmatrix}$. Where does ∇g vanish? This is the pair of simultaneous conditions $2x - y = 0$ and $-x + 2y = 0$, the only solution to which is $(0, 0)$ (this is "2 equations in 2 unknowns" as considered in high school algebra). Since $(0, 0)$ isn't on the level set $g(x, y) = 3$ (since $g(0, 0) = 0$), there are no bad points to worry about. So at any local extremum (x, y) for f on $g = 3$ we must have $(\nabla f)(x, y) = \lambda(\nabla g)(x, y)$ for some scalar λ, which is to say

$$\begin{bmatrix} y \\ x \end{bmatrix} = \lambda \begin{bmatrix} 2x - y \\ -x + 2y \end{bmatrix}.$$

Hence, we want to find all solutions to the combined system of equations

$$y = \lambda(2x - y), \;\; x = \lambda(-x + 2y), \;\; 3 = x^2 - xy + y^2$$

(where λ is also unknown, and we don't really care so much about its value); the final condition in this combined list of three equations is what records that we are finding a point on the level set $g = 3$ of interest.

We will use the first two equations in our combined system to get two different expressions for λ that we then compare to get another condition on x and y that we exploit in conjunction with the third equation. But please always remember: *never divide by* 0. In other words, the first two equations in the combined system give rise to two expressions for λ as

$$\frac{y}{2x - y} = \lambda = \frac{x}{-x + 2y},$$

but to write these requires knowing that the denominators don't vanish: $2x - y \neq 0$ and $-x + 2y \neq 0$.

So let's first handle the problematic cases where $2x - y = 0$ **or** $-x + 2y = 0$ (this is not "and", but "or": make sure you understand the distinction, and why it matters for a systematic analysis of all possibilities!). If $2x - y = 0$ then the first equation forces $y = 0$ (regardless of what λ is), but then the condition $2x - y = 0$ forces $x = 0$ too, and we've already discussed that $(0, 0)$ isn't on the level set (i.e., it violates the third equation in our combined system).

Hence, for our purposes we can assume $2x - y \neq 0$. The exact same reasoning works in the case $-x + 2y = 0$ via the second equation in our combined system. Hence, we can also assume $-x + 2y \neq 0$, so our two expressions for λ are both valid. Equating them says

$$\frac{y}{2x - y} = \frac{x}{-x + 2y},$$

and cross-multiplying gives $y(-x+2y) = x(2x-y)$, or equivalently $-xy+2y^2 = 2x^2-xy$, which says $2x^2 = 2y^2$, or in other words $x = \pm y$.

Copyright © 2018 Stanford University Department of Mathematics. All rights reserved.

Hence, our conditions are $x = \pm y$ and $3 = x^2 - xy + y^2$ (and then we can solve for λ using our expressions for it, but we actually don't care what λ is for our purposes). Thus, we separately treat each of the cases $x = y$ and $x = -y$ along with the equation $3 = x^2 - xy + y^2$ that is the "level set" condition.

If $x = y$ then $3 = x^2 - xy + y^2 = y^2 - y^2 + y^2 = y^2$, which says $y = \pm\sqrt{3}$. Since $x = y$, we arrive at the candidates $(\sqrt{3}, \sqrt{3})$ and $(-\sqrt{3}, -\sqrt{3})$.

Next, if $x = -y$ then $3 = (-y)^2 - (-y)y + y^2 = y^2 + y^2 + y^2 = 3y^2$, or equivalently $1 = y^2$, so $y = \pm 1$ and then $x = -y = \mp 1$. In other words, we get two more candidates: $(1, -1)$ and $(-1, 1)$.

Finally, we compute the function $f(x, y) = xy$ at each of these 4 candidates: $f(\sqrt{3}, \sqrt{3}) = 3$, $f(-\sqrt{3}, -\sqrt{3}) = 3$, $f(1, -1) = -1$ and $f(-1, 1) = -1$. Inspecting which of these values are biggest and smallest, we conclude that the maximum for xy on $g = 3$ is 3, attained at the points $(\sqrt{3}, \sqrt{3})$ and $(-\sqrt{3}, -\sqrt{3})$, and the minimum for xy on $g = 3$ is -1, attained at the points $(1, -1)$ and $(-1, 1)$.

(b) Definitely less appetizing!

Problem 2: Optimization review (what technique(s) would you use?)

(a) Given the function $f(x, y) = x + y$, find the maximum and minimum values of f on the domain

$$D_1 = \{(x, y) \colon 0 \le y \le x^2 \text{ and } -1 \le x \le 1\}.$$

(b) Find the maximum and minimum values of $G(x, y) = 3x^2 + 4xy$ on the region

$$D_2 = \{(x, y) \colon y \ge 0 \text{ and } x^2 + y^2 \le 9\}.$$

(When doing this, one part of the boundary will be a mess via single-variable calculus, so employ Lagrange multipliers there with the boundary curve as a constraint condition. You may encounter the expression $2x^2 - 3xy - 2y^2$, in which case it will be useful to then observe that this factors as $(2x + y)(x - 2y)$.)

(c) (**Extra**) Let C be the curve in $\mathbf{R}^2$ defined by the equation

$$y^2 = x^3 - 4x^2 + 5x$$

Determine all points on C at minimal distance to $(5/2, 0)$.

Solution:

(a) We'll search the following regions:

$$\begin{array}{ccccc} \text{interior} & \text{top edge} & \text{bottom edge} & \text{left edge} & \text{right edge} \\ \begin{bmatrix} 0 < y < x^2 \\ -1 < x < 1 \end{bmatrix} & \begin{bmatrix} y = x^2 \\ -1 \le x \le 1 \end{bmatrix} & \begin{bmatrix} y = 0 \\ -1 \le x \le 1 \end{bmatrix} & \begin{bmatrix} x = -1 \\ 0 \le y \le 1 \end{bmatrix} & \begin{bmatrix} x = 1 \\ 0 \le y \le 1 \end{bmatrix} \end{array}$$

interior. First we compute that $\nabla f = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \neq \mathbf{0}$, so f has no critical points in the interior.

For potential extrema on the boundary, we use single-variable calculus and treat the boundary parts one at a time.

top edge. Here we have $f = x + x^2$ across the interval $-1 \le x \le 1$. We have $\frac{d}{dx}(x + x^2) = 1 + 2x$, which equals zero when $x = -1/2$, a value in the interval in question. Thus, this point $(x, y) = (-1/2, 1/4)$ and the endpoints $(-1, 1)$ and $(1, 1)$ are put on the list of candidates for where the extrema may be attained.

bottom edge. Here we have $f = x$ across the interval $-1 \le x \le 1$. Although $\frac{d}{dx}(x) = 1$ never vanishes, we must put the endpoints $(-1, 0)$ and $(1, 0)$ on the list of candidates.

Copyright © 2018 Stanford University Department of Mathematics. All rights reserved.

left edge. Here we have $f = -1 + y$ across the interval $\{0 \leq y \leq 1\}$. Since the single-variable derivative of this expression never vanishes, we just need to include the endpoints $(-1, 0)$ and $(-1, 1)$ among the candidates.

right edge. Here we have $f = 1 + y$ across the interval $\{0 \leq y \leq 1\}$. As before, therefore we just need to put the endpoints $(1, 0)$ and $(1, 1)$ are among the candidates.

Our overall list of candidates for locations of where f attains extreme values on D_1 are: $(-1/2, 1/4)$, $(-1, 1)$, $(1, 1)$, $(-1, 0)$, $(1, 0)$. By computing the value of f at each of these points, we see that the maximum value is $f(1, 1) = 2$ and the minimum value is $f(-1, 0) = -1$.

(b) We'll search the following regions for candidate extrema:

$$\begin{array}{ccc} \text{interior} & \text{bottom edge} & \text{top semicircle} \\ \begin{bmatrix} x^2 + y^2 < 9 \\ y > 0 \end{bmatrix} & \begin{bmatrix} y = 0 \\ -3 \leq x \leq 3 \end{bmatrix} & \begin{bmatrix} x^2 + y^2 = 9 \\ y \geq 0 \end{bmatrix} \end{array}$$

interior First we compute that $\nabla G = \begin{bmatrix} 6x + 4y \\ 4x \end{bmatrix}$, which equals $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ only for $(x, y) = (0, 0)$, which lies in D_2, so it has to be on the list of candidates where extreme values might be attained. (This critical point is not actually in the interior — it's on the bottom boundary — so we will see it crop up as a potential extremum along that edge below. But it never hurts to add this point to the list now!)

For potential extrema on the boundary, we use single-variable calculus and treat the boundary parts (bottom edge and top semicircle) one at a time.

bottom edge. Here we have $G = 3x^2$ across the interval $-3 \leq x \leq 3$. Since $\frac{d}{dx}(3x^2) = 6x$ vanishes when $x = 0$, which lies in this interval, we have three points to include in the list of candidates: $(0, 0)$ and the endpoints $(-3, 0)$ and $(3, 0)$.

top semicircle. This is a circular arc rather than a line segment as for the bottom boundary, and trying to use single-variable calculus leads to a mess for the resulting single-variable expression for G. So away from the endpoints of this semicircle we'll employ Lagrange multipliers by viewing that curve as a constraint condition! Here the constraint is $g = x^2 + y^2 = 9$, for which $\nabla g = \begin{bmatrix} 2x \\ 2y \end{bmatrix}$, which never vanishes on the constraint region (it only vanishes at the origin, which is not on $g = 9$).

Thus, there are no "problematic" points for the Lagrange multiplier method and hence we seek points on this semicircle away from its endpoints where $\nabla G = \lambda(\nabla g)$ for some λ. This condition, along with the constraint condition $g = 9$, is expressed by the combined system of equations

$$6x + 4y = 2\lambda x, \;\; 4x = 2\lambda y, \;\; x^2 + y^2 = 9.$$

As usual, we "solve for λ" in each of the first two equations and set the resulting expressions equal to each other to get a new condition on x and y (that can be combined with the constraint equation). But as always, we must be careful about cases that would lead to division by 0. So we first "solve for λ" in each of the first two equations in the combined system to figure out what are the potentially problematic division-by-zero cases:

$$\frac{6x + 4y}{2x} = \lambda = \frac{4x}{2y}.$$

Hence, we first deal with situations where either $x = 0$ or $y = 0$. If $x = 0$ then the first equation in our combined system says $0 + 4y = 0$ (regardless of λ), so $y = 0$, but we have already noted that $(0, 0)$ violates the constraint condition $g = 9$, so this possibility cannot occur. Likewise, if $y = 0$ then the second equation in our combined system says $4x = 0$ (regardless of λ), forcing $x = 0$, so again we are at the point $(0, 0)$ which has aleady been ruled out. So there is no division-by-zero problem.

Coming back to our two expressions for λ, setting them equal to each other says

$$\frac{6x + 4y}{2x} = \frac{4x}{2y},$$

Copyright © 2018 Stanford University Department of Mathematics. All rights reserved.

and cross-multiplying converts this into the equation

$$2y(6x + 4y) = (2x)(4x),$$

or equivalently $12xy + 8y^2 = 8x^2$, which is the same as $3xy + 2y^2 = 2x^2$. Remembering the constraint equation, we have arrived at two equations on x and y without reference to λ:

$$x^2 + y^2 = 9, \;\; 2x^2 - 3xy - 2y^2 = 0$$

with $x, y \neq 0$. Ah, but we were given the hint to use that $2x^2 - 3xy - 2y^2 = (2x + y)(x - 2y)$, whose vanishing happens precisely when $x = 2y$ or $y = -2x$.

So we now treat the possibilities $x = 2y$ or $y = -2x$ separately. If $x = 2y$ then the constraint $x^2 + y^2 = 9$ says $(2y)^2 + y^2 = 9$, or equivalently $5y^2 = 9$, so $y = \pm 3/\sqrt{5}$ and then $x = 2y = \pm 6/\sqrt{5}$ (same sign as for y). If instead $y = -2x$ then $9 = x^2 + y^2 = x^2 + (-2x)^2 = 5x^2$, so $x = \pm 3/\sqrt{5}$ and then $y = -2x = \mp 6/\sqrt{5}$. But recall that we are on the upper semicircle with $y \geq 0$, so we're left with just two points obtained in this way: $(6/\sqrt{5}, 3/\sqrt{5})$ and $(-3/\sqrt{5}, 6/\sqrt{5})$.

Thus, our final list of candidate points for extrema of $G(x, y) = 3x^2 + 4xy$ on D_2 is:

$$(0, 0), \;\; (-3, 0), \;\; (3, 0), \;\; (6/\sqrt{5}, 3/\sqrt{5}), \;\; (-3/\sqrt{5}, 6/\sqrt{5}).$$

We have $G(0, 0) = 0$, $G(-3, 0) = 27$, $G(3, 0) = 27$, $G(6/\sqrt{5}, 3/\sqrt{5}) = 36$, and $G(-3/\sqrt{5}, 6/\sqrt{5}) = 27/5 - 72/5 = -45/5 = -9$. Hence, the maximal value of G on this region is 36 (attained only at $(6/\sqrt{5}, 3/\sqrt{5})$) and the minimal value is -9 (attained only at $(-3/\sqrt{5}, 6/\sqrt{5})$).

(c) We work with the squared distance function to $(5/2, 0)$, namely $f(x, y) = (x - 5/2)^2 + y^2$. The Lagrange Multiplier Theorem states that for $g(x, y) = y^2 - x^3 + 4x^2 - 5x$, the points where the function $f(x, y)$ is minimized on the level set $C = \{g(x, y) = 0\}$ are among the points where either $\nabla f = \lambda \nabla g$, for some real number λ, or where $\nabla g = \mathbf{0}$.

The vanishing of $\nabla g = \begin{bmatrix} -3x^2 + 8x - 5 \\ 2y \end{bmatrix}$ happens precisely when $y = 0$ and $-3x^2 + 8x - 5 = 0$. By the quadratic formula, the latter condition on x has as its solutions exactly $x = 1, 5/3$, so this singles out the points $(1, 0)$ and $(5/3, 0)$ for separate treatment. But these points can't occur because neither of them satisfies the constraint condition $g(x, y) = 0$ (indeed, $g(1, 0) = -1 + 4 - 5 = -2 \neq 0$ and $g(5/3, 0) = -125/27 + 400/9 - 25/3 = 850/3 \neq 0$).

So now we may focus on the multiplier condition $\nabla f = \lambda(\nabla g)$, which says

$$\begin{bmatrix} 2x - 5 \\ 2y \end{bmatrix} = \lambda \begin{bmatrix} -3x^2 + 8x - 5 \\ 2y \end{bmatrix}.$$

Expressing this as a pair of scalar equations and bringing in the constraint equation $g = 0$, we arrive at the combined system of three equations

$$2x - 5 = \lambda(-3x^2 + 8x - 5), \;\; 2y = \lambda(2y), \;\; y^2 - x^3 + 4x^2 - 5x = 0.$$

As usual, we "solve for λ" in each of the first two equations, assuming denominators don't vanish: we obtain

$$\frac{2x - 5}{-3x^2 + 8x - 5} = \lambda = \frac{2y}{2y}.$$

This works as long as both denominators are nonzero, so first we dispoose of the problematic case where some denominator vanishes: either $y = 0$ or $-3x^2 + 8x - 5$.

Case 1: Suppose $y = 0$. The second multiplier equation then tells us nothing, and we also learn nothing from the first: is has nothing to do with y (and we haven't learned anything about λ yet). But going back to the constraint equation with $y = 0$, we have $0 = g(x, 0) = -x^3 + 4x^2 - 5x = -x(x^2 - 4x + 5)$. $x^2 - 4x + 5 = (x - 2)^2 + 1 = 0$ has no real solutions, so $x = 0$ is the only solution. In other words, we have singled out the point $(0, 0)$ as needing separate treatment. We put this into the fridge and will come back to them later.

Case 2: Suppose $-3x^2 + 8x - 5 = 0$, which (by the quadratic formula) says $x = 1, 5/3$. The first multiplier equation then gives $2x - 5 = 0$ (regardless of λ), so $x = 5/2$; this is absurd since we are only allowing $x = 1, 5/3$, so this situation cannot occur.

Copyright © 2018 Stanford University Department of Mathematics. All rights reserved.

Now returning to the original combined system (away from the problematic points already identified), we have (with non-vanishing denominators) the two fraction expressions for λ given above. Setting them equal to each other and cross-multiplying gives

$$2y(2x-5) = 2y(-3x^2+8x-5).$$

But we have $y \neq 0$ (since we are avoiding the problematic points for now), so cancelling $2y \neq 0$ from both sides gives $2x-5 = -3x^2+8x-5$, or equivalently $3x^2-6x=0$, which says $3x(x-2)=0$, so $x=0$ or $x=2$. Then the constraint equation $g(x,y)=0$ says $y^2 = x^3-4x^2+5x$ is equal to 0 when $x=0$ and is equal to $8-16+10=2$ when $x=2$. In other words, if $x=0$ then $y=0$ and if $x=2$ then $y=\pm\sqrt{2}$. Hence, we obtain the further candidate points $(0,0)$ and $(2,\pm\sqrt{2})$.

Putting it all together, all local minima for the squared distance $f(x,y)$ on the constraint curve $g=0$ must occur among the following points on the constraint curve: $(0,0)$, $(2,\pm\sqrt{2})$. Evaluating $f(x,y)=(x-5/2)^2+y^2$ at these points gives the values $f(0,0)=25/4$, $f(2,\pm\sqrt{2})=1/4+2=9/4$. The least value among these is $9/4$, attained only at the points $(2,\pm\sqrt{2})$, so these two points on C attain the least distance to $(5/2,0)$, with minimal distance equal to $\sqrt{9/4}=3/2$.

Problem 3: Identifying linear functions

In each case below, is $\mathbf{f} : \mathbf{R}^2 \to \mathbf{R}^3$ linear? If it is, find the matrix representing it. If not, explain why not.

(a) $\mathbf{f}(x_1,x_2) = (x_1, x_2^2, 2x_1+x_2)$

(b) $\mathbf{f}(x_1,x_2) = (1, x_2, 2x_1+x_2)$

(c) $\mathbf{f}(x_1,x_2) = (0, x_2, 2x_1+x_2)$

(d) $\mathbf{f}(x_1,x_2) = (0, x_1x_2, 2x_1+x_2)$

(e) $\mathbf{f}(x_1,x_2) = (ax_1+bx_2, cx_1+dx_2, ex_1+fx_2)$

Solution: (c) is linear, coming from the matrix $\begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 2 & 1 \end{bmatrix}$; (e) is linear, coming from the matrix $\begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}$.

The rest are not linear, since some vector entry in the output is not a linear combination of the input variables and so cannot arise as a matrix-vector product against $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

Problem 4: Derivative matrix and numerical linear approximation

Consider the function $f : \mathbf{R}^2 \to \mathbf{R}^2$ given by

$$f(x,y) = (x^3y^2, 4x+y^3+xy).$$

(a) Compute the derivative matrix $(Df)(x,y)$, and then use it to give the linear approximation to f at $(1,1)$.

(b) Use your answer to (a) to estimate the 2-vector $f(0.8,1.1)$, and then compare it with an exact calculation using a calculator. Is it a good approximation?

(c) Give the linear approximation to f at $(2,-2)$ and use it to estimate the 2-vector $f(2.1,-1.9)$ and then compare this to the exact 2-vector using a calculator. Is the approximation good or bad?

Copyright © 2018 Stanford University Department of Mathematics. All rights reserved.

Solution:

(a) By computing partial derivatives of the component functions of f, we have $(Df)(x, y) = \begin{bmatrix} 3x^2y^2 & 2x^3y \\ 4+y & 3y^2+x \end{bmatrix}$, so $(Df)(1, 1) = \begin{bmatrix} 3 & 2 \\ 5 & 4 \end{bmatrix}$. Hence, for s, t near 0 we have

$$f(1+s, 1+t) \approx f(1, 1) + \begin{bmatrix} 3 & 2 \\ 5 & 4 \end{bmatrix} \begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \end{bmatrix} + \begin{bmatrix} 3s+2t \\ 5s+4t \end{bmatrix} = \begin{bmatrix} 1+3s+2t \\ 6+5s+4t \end{bmatrix}.$$

(b) We set $s = -0.2$ and $t = 0.1$ in the linear approximation in (a) to get the estimate

$$f(0.8, 1.1) \approx \begin{bmatrix} 1-0.6+0.2 \\ 6-1.0+0.4 \end{bmatrix} = \begin{bmatrix} 0.6 \\ 5.4 \end{bmatrix}.$$

An exact calculation with a calculator gives that $f(0.8, 1.1) = \begin{bmatrix} .61952 \\ 5.411 \end{bmatrix}$, so the approximation turned out quite well!

(c) By (a), $(Df)(2, -2) = \begin{bmatrix} 48 & -32 \\ 2 & 14 \end{bmatrix}$, so the linear approximation at $(2, -2)$ says that for s, t near 0 we have

$$f(2+s, -2+t) \approx f(2, -2) + \begin{bmatrix} 48 & -32 \\ 2 & 14 \end{bmatrix} \begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} 32 \\ -4 \end{bmatrix} + \begin{bmatrix} 48s-32t \\ 2s+14t \end{bmatrix} = \begin{bmatrix} 32+48s-32t \\ -4+2s+14t \end{bmatrix}.$$

Hence, with $s = 0.1$ and $t = 0.1$ we get

$$f(2.1, -1.9) \approx \begin{bmatrix} 32+48(0.1)-32(0.1) \\ -4+2(0.1)+14(0.1) \end{bmatrix} = \begin{bmatrix} 32+4.8-3.2 \\ -4+0.2+1.4 \end{bmatrix} = \begin{bmatrix} 33.6 \\ -2.4 \end{bmatrix}.$$

Using a calculator, we have the exact result $f(2.1, -1.9) = (33.43221, -2.449)$. So in this case the linear approximation is still not bad.

Copyright © 2018 Stanford University Department of Mathematics. All rights reserved.