LINEAR ALGEBRA AND VECTOR ANALYSIS

MATH 22A

Unit 20: Constraints

Lecture

20.1. If we want to maximize a function $f : \mathbb{R}^m \to \mathbb{R}$ on the constraint $S = \{x \in \mathbb{R}^m \mid g(x) = c\}$, then both the gradients of f and g matter. We call two vectors v, w **parallel** if $v = \lambda w$ or $w = \lambda v$ for some real λ. The zero vector is parallel to everything. Here is a variant of Fermat:

> **Theorem:** If x_0 is a maximum of f under the constraint $g = c$, then $\nabla f(x_0)$ and $\nabla g(x_0)$ are parallel.

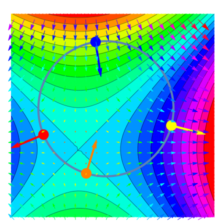
20.2. Proof: use contradiction: assume $\nabla f(x_0)$ and $\nabla g(x_0)$ are not parallel and x_0 is a local maximum. Let T be the tangent plane to $S = \{g = c\}$ at x_0. Because $\nabla f(x_0)$ is not perpendicular to T we can project it onto T to get a non-zero vector v in T which is not perpendicular to ∇f. Actually the angle between ∇f and v is acute so that $\cos(\alpha) > 0$. Take a curve $r(t)$ in S with $r(0) = x_0$ and $r'(0) = v$. We have $d/dt f(r(0)) = \nabla f(r(0)) \cdot r'(0) = |\nabla f(x_0)||v| \cos(\alpha) > 0$. By linear approximation, we know that $f(r(t)) > f(r(0))$ for small enough $t > 0$. This is a contradiction to the fact that f was maximal at $x_0 = r(0)$ on S.

20.3. This immediately implies: (distinguish $\nabla g \neq 0$ and $\nabla g = 0$)

> **Theorem:** For a maximum of f on $S = \{g = c\}$ either the Lagrange equations $\nabla f(x_0) = \lambda \nabla g(x_0), g = c$ hold, or then $\nabla g(x_0) = 0, g = c$.

20.4. For functions $f(x, y), g(x, y)$ of two variables, this means we have to solve a system with three equations and three unknowns:

$$\begin{array}{rcl} f_x(x_0, y_0) & = & \lambda g_x(x_0, y_0) \\ f_y(x_0, y_0) & = & \lambda g_y(x_0, y_0) \\ g(x, y) & = & c \end{array}$$



20.5. To find a maximum, solve the Lagrange equations and add a list of critical points of g on the constraint. Then pick a point where f is maximal among all points. We don't bother with a second derivative test. But here is a possible statement:

$$\frac{d^2}{dt^2} D_{tv} D_{tv} f(x_0)|_{t=0} < 0$$

for all v perpendicular to $\nabla g(x_0)$, then x_0 is a local maximum.

20.6. Of course, the case of maxima and minima are analog. If f has a maximum on $g = c$, then $-f$ has a minimum at $g = c$. We can have a maximum of f under a smooth constraint $S = \{g = c\}$ without that the Lagrange equations are satisfied. An example is $f(x, y) = x$ and $g(x, y) = x^3 - y^2$ shown in Figure (1).

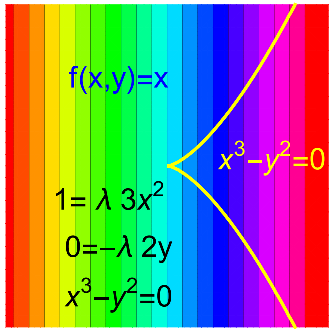


FIGURE 1. An example of a function, where the Lagrange equations do not give the minimum, here $(0, 0)$. It is a case, where $\nabla g = 0$.

20.7. The method of Lagrange can maximize functions f under several constraints. Lets show this in the case of a function $f(x, y, z)$ of three variables and two constraints $g(x, y, z) = c$ and $h(x, y, z) = d$. The analogue of the Fermat principle is that at a maximum of f, the gradient of f is in the plane spanned by ∇g and ∇h. This leads to the **Lagrange equations** for 5 unknowns x, y, z, λ, μ.

$$\begin{array}{rcl} f_x(x_0, y_0, z_0) & = & \lambda g_x(x_0, y_0, z_0) + \mu h_x(x_0, y_0, z_0) \\ f_y(x_0, y_0, z_0) & = & \lambda g_y(x_0, y_0, z_0) + \mu h_y(x_0, y_0, z_0) \\ f_z(x_0, y_0, z_0) & = & \lambda g_z(x_0, y_0, z_0) + \mu h_z(x_0, y_0, z_0) \\ g(x, y, z) & = & c \\ h(x, y, z) & = & d \end{array}$$

20.8. For example, if $f(x, y, z) = x^2 + y^2 + z^2$ and $g(x, y, z) = x^2 + y^2 = 1$, $h(x, y, z) = x + y + z = 4$, then we find points on the ellipse $g = 1, h = 4$ with minimal or maximal distance to 0.

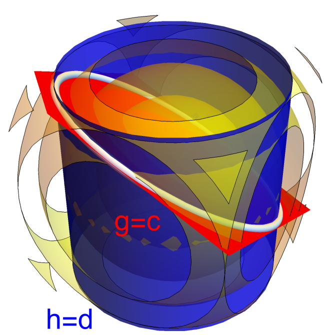


FIGURE 2. Extremizing a function f under two constraints. In this case the intersection $g = c, h = d$ is an ellipse.

Examples

20.9. Problem: Minimize $f(x,y) = x^2+2y^2$ under the constraint $g(x,y) = x+y^2 = 1$. **Solution:** The Lagrange equations are $2x = \lambda, 4y = \lambda 2y$. If $y = 0$ then $x = 1$. If $y \neq 0$ we can divide the second equation by y and get $2x = \lambda, 4 = \lambda 2$ again showing $x = 1$. The point $x = 1, y = 0$ is the only solution.

20.10. Problem: Which cylindrical soda can of height h and radius r has minimal surface A for fixed volume V? **Solution:** We have $V(r,h) = h\pi r^2 = 1$ and $A(r,h) = 2\pi rh + 2\pi r^2$. With $x = h\pi, y = r$, you need to optimize $f(x,y) = 2xy + 2\pi y^2$ under the constrained $g(x,y) = xy^2 = 1$. We will do that in class.

20.11. Problem: If $0 \leq p_k \leq 1$ is the probability that a dice shows k, then we have $g(p) = p_1 + p_2 + \cdots + p_6 = 1$. This vector p is called a **probability distribution**. The **Shannon entropy** of p is defined as

$$S(p) = -\sum_{i=1}^{6} p_i \log(p_i) = -p_1 \log(p_1) - p_2 \log(p_2) - \cdots - p_6 \log(p_6) .$$

Find the distribution p which maximizes entropy S. **Solution:** $\nabla f = (-1 - \log(p_1), \dots, -1 - \log(p_n))$, $\nabla g = (1, \dots, 1)$. The Lagrange equations are $-1 - \log(p_i) = \lambda, p_1 + \cdots + p_6 = 1$, from which we get $p_i = e^{-(\lambda+1)}$. The last equation $1 = \sum_i \exp(-(\lambda+1)) = 6\exp(-(\lambda+1))$ fixes $\lambda = -\log(1/6) - 1$ so that $p_1 = p_2 = \cdots = p_6 = 1/6$. It is the fair dice that has maximal entropy. Maximal entropy means **least information content**.

20.12. Assume that the probability that a physical or chemical system is in a state k is p_k and that the energy of the state k is E_k. Nature minimizes the **free energy**

$$F(p_1, \dots, p_n) = -\sum_i [p_i \log(p_i) - E_i p_i]$$

if the energies E_i are fixed. The probability distribution p_i satisfying $\sum_i p_i = 1$ minimizing the free energy is called a **Gibbs distribution**. Find this distribution in general if E_i are given. **Solution:** $\nabla f = (-1 - \log(p_1) - E_1, \dots, -1 - \log(p_n) - E_n)$, $\nabla g = (1, \dots, 1)$. The Lagrange equation are $\log(p_i) = -1 - \lambda - E_i$, or $p_i = \exp(-E_i)C$, where $C = \exp(-1-\lambda)$. The constraint $p_1 + \cdots + p_n = 1$ gives $C(\sum_i \exp(-E_i)) = 1$ so that $C = 1/(\sum_i e^{-E_i})$. The **Gibbs solution** is $p_k = \exp(-E_k)/\sum_i \exp(-E_i)$. [1]

20.13. If f is a quadratic function on $\mathbb{R}^m$ and g is linear that is $f(x) = Bx \cdot x/2$ with $B \in M(m,m)$ and if the constraint $g(x) = Ax = c$ is linear $A \in M(1,m)$, then $\nabla f(x) = Bx$ and $\nabla g(x) = A^T$. Lets call $b = A^T \in M(m,1) \sim \mathbb{R}^m$. The Lagrange equations are then $Bx = \lambda b, Ax = c$. We see in general that for quadratic f and linear g, we end up with a **linear system of equations.**

20.14. Related to the previous remark is the following observation. It is often possible to reduce the Lagrange problem to a problem without constraint. This is a point of view often taken by economists. Let us look at it in dimension 2, where we extremize $f(x,y)$ under the constraint $g(x,y) = 0$. Define $F(x,y,\lambda) = f(x,y) - \lambda g(x,y)$. The Lagrange equations for f, g are now equivalent to $\nabla F(x,y,\lambda) = 0$ in three dimensions.

[1] This example is from Rufus Bowen, Lecture Notes in Math, 470, 1978

Homework

Problem 20.1: Find the cylindrical basket which is open on the top has has the largest volume for fixed area π. If x is the radius and y is the height, we have to maximize $f(x,y) = \pi x^2 y$ under the constraint $g(x,y) = 2\pi xy + \pi x^2 = \pi$. Use the method of Lagrange multipliers.

Problem 20.2: Given a $n \times n$ symmetric matrix B, we look at the function $f(x) = x \cdot Bx$. and look at extrema of f under the constraint that $g(x) = x \cdot x = 1$. This leads to an equation

$$Bx = \lambda x \ .$$

A solution x is called an **eigenvector**. The Lagrange constant λ is an **eigenvalue**. Find the solutions to $Bx = \lambda x, |x| = 1$ if B is a 2×2 matrix, where $f(x,y) = ax^2 + (b+c)xy + dy^2$ and $g(x,y) = x^2 + y^2$. Then solve the problem where $a = 3, b = 2, c = 4, d = 1$. (Never mind here that B is not symmetric).

Problem 20.3: Which pyramid of height h over a square $[-a,a] \times [-a,a]$ with surface area is $4a\sqrt{h^2 + a^2} + 4a^2 = 4$ has maximal volume $V(h,a) = 4ha^2/3$? By using new variables (x,y) and multiplying V with a constant, we get to the equivalent problem to maximize $f(x,y) = yx^2$ over the constraint $g(x,y) = x\sqrt{y^2 + x^2} + x^2 = 1$. Use the later variables.

Problem 20.4: Motivated by the Disney movie "Tangled", we want to build a hot air balloon with a cuboid mesh of dimension x, y, z which together with the top and bottom fortifications uses wires of total length $g(x,y,z) = 6x + 6y + 4z = 32$. Find the balloon with maximal volume $f(x,y,z) = xyz$.

Problem 20.5: A **solid bullet** made of a half sphere and a cylinder has the volume $V = 2\pi r^3/3 + \pi r^2 h$ and surface area $A = 2\pi r^2 + 2\pi rh + \pi r^2$. Doctor Manhattan designs a bullet with fixed volume and minimal area. With $g = 3V/\pi = 1$ and $f = A/\pi$ he therefore minimizes $f(h,r) = 3r^2 + 2rh$ under the constraint $g(h,r) = 2r^3 + 3r^2 h = 1$. Use the Lagrange method to find a local minimum of f under the constraint $g = 1$.

Oliver Knill, knill@math.harvard.edu, Math 22a, Harvard College, Fall 2018