Optimization – Part 2

Contents

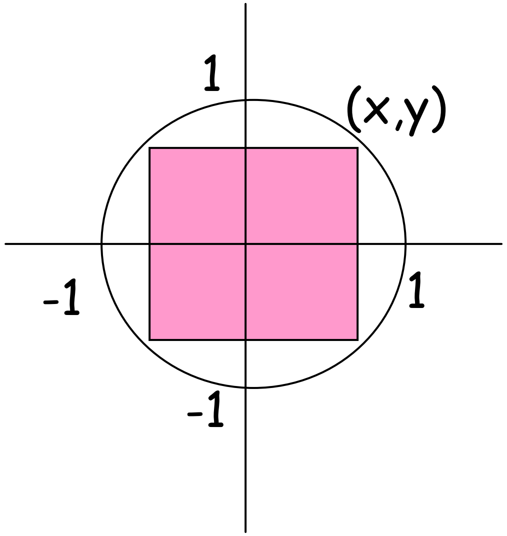
- Terms and definitions, formulating the problem
- Unconstrained optimization conditions
  - FONC, SONC, SOSC
- Searching for the solution
  - Which direction?
  - What step size?
- Constrained optimization
- Lagrangian Multipliers, the Lagrangian function
- KKT conditions
- The Dual

Constrained optimization

$$
\begin{array}{lll}
\min_x & f(x) & \\
\text{subject to} & g(x) \leq 0 & \text{Inequality constraints} \\
 & h(x) = 0 & \text{Equality constraints} \\
 & x \in R &
\end{array}
$$

Example 1: Nonlinear equality constraints

What is the largest rectangle (areawise) that can be drawn inside a circle of radius 1?



Let (x,y) be the coordinates of the top right corner.

The area of the square is 4xy.

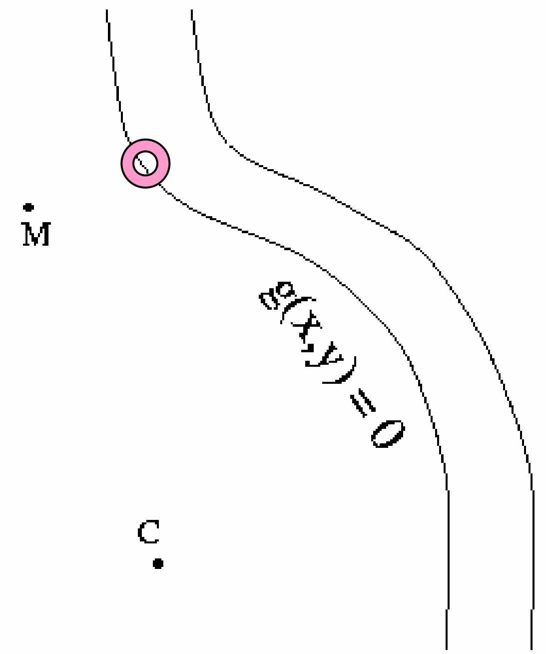
So we wish to maximize 4xy subject to the constraint $x^2 + y^2 = 1$

$$\begin{array}{ll} \max & 4xy \\ \text{subject to} & x^2 + y^2 \leq 1 \end{array}$$

Lagrangian Multipliers …

A classic example: the "milkmaid problem"

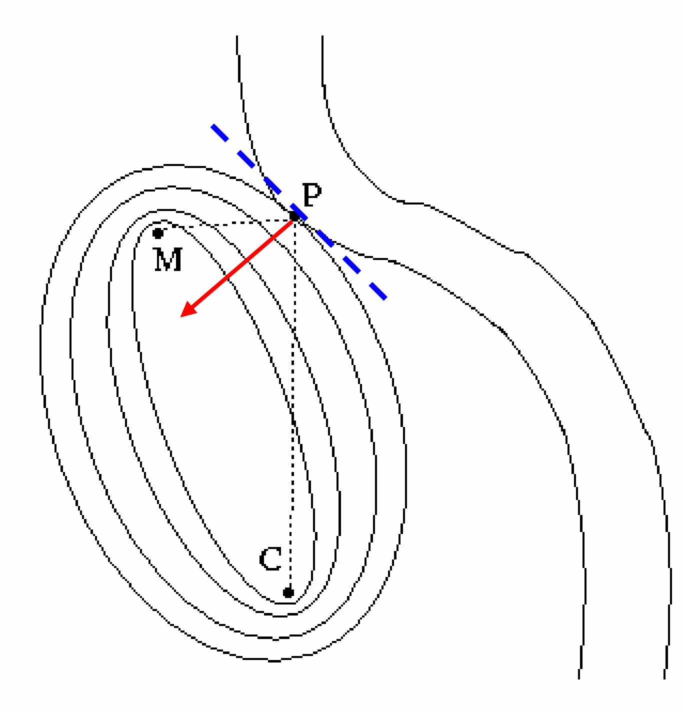
It's milking time at the farm, and the milkmaid has been sent to the field to get the day's milk. She's in a hurry to get back for a date with a handsome young goatherd, so she wants to finish her job as quickly as possible. However, before she can gather the milk, she has to rinse out her bucket in the nearby river.
Just when she reaches point **M**, our heroine spots the cow, way down at point **C**. Because she is in a hurry, she wants to take the shortest possible path from where she is to the river and then to the cow. If the near bank of the river is a curve satisfying the function $g(x,y) = 0$, what is the shortest path for the milkmaid to take? (To keep things simple, we assume that the field is flat and uniform and that all points on the river bank are equally good.)
To put this into more mathematical terms, the milkmaid wants to find the point **P** for which the distance d(**M**,**P**) from **M** to **P** plus the distance d(**P**,**C**) from **P** to **C** is a minimum. We have to impose the *constraint* that **P** is a point on the riverbank. Formally, we must minimize the function $f(P) = d(M,P) + d(P, C)$, subject to the constraint that $g(P) = 0$.



http://www.slimy.com/~steuard/teaching/tutorials/Lagrange.html

Lagrangian Multipliers … aka the **dual variables**

Geometry Fact: For every point **P** on a given ellipse, the total distance from one focus of the ellipse to **P** and then to the other focus is exactly the same. In our problem, that means that the milkmaid could get to the cow by way of any point on a given ellipse in the same amount of time. Therefore, we just have to find the smallest ellipse that intersects the curve of the river. The image shows a sequence of ellipses of larger and larger size whose foci are **M** and **C**, ending with the one that is just *tangent* to the riverbank … at the ideal point **P**.
Tangent → Same gradient. Therefore at P, the **normal vector** to the ellipse is in the same direction as the normal vector to the riverbank. Mathematically,



$$\nabla f(P) = \lambda \, \nabla g(P)$$

Lagrangian multipliers express the gradient at the optimum as a linear combination of the constraints.

The Lagrangian Function

Equality constraints (just like the milkmaid problem)

$$\begin{array}{ll} \min_x & f(x) \\ \text{subject to} & h_i(x) = 0 \end{array}$$

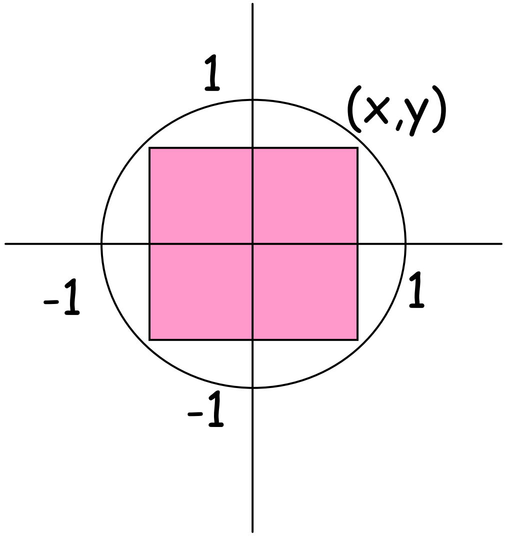
The Lagrangian function is defined to be:

$$L(x, \lambda) = f(x) - \sum_i \lambda_i h_i(x)$$

And from our conclusion on Lagrangian multipliers, at x* the following is true.

$$\bigtriangledown L(x, \lambda) = \bigtriangledown f(x) - \bigtriangledown \sum_i \lambda_i h_i(x) = 0$$

Solving Example 1



$$\begin{array}{ll} \max & 4xy \\ \text{subject to} & x^2 + y^2 \leq 1 \end{array}$$

The Lagrangian function is …

$$L(x, y, \lambda) = 4xy - \lambda(x^2 + y^2 - 1)$$

At the optimal solution, the derivative of the Lagrangian function is …

$$\bigtriangledown L(x, y, \lambda) = \bigtriangledown 4xy - \bigtriangledown \lambda(x^2 + y^2 - 1) = 0$$

$$\nabla L(x, y, \lambda) = \nabla 4xy - \nabla \lambda(x^2 + y^2 - 1) = 0$$

$$\nabla L = \begin{pmatrix} 4y \\ 4x \end{pmatrix} - \begin{pmatrix} 2\lambda x \\ 2\lambda y \end{pmatrix} = 0$$

Therefore, at the optimal solution, the following equations hold:

$$x^2 + y^2 \leq 1$$
$$4y - 2\lambda x = 0$$
$$4x - 2\lambda y = 0$$

Solutions:

When $\lambda = 2, \pm(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ max

When $\lambda = -2, \pm(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$ min

KKT conditions

$$\begin{array}{ll} \min_x & f(x) \\ \text{subject to} & g(x) \leq 0 \\ & h(x) = 0 \end{array}$$

$$L(x, \lambda, \mu) = f(x) - \lambda g(x) - \mu h(x)$$

The KKT conditions for x* are as follows:

$$\begin{array}{l} \triangledown f(x^*) = \lambda \triangledown g(x^*) + \mu \triangledown h(x^*) \\ \lambda g(x^*) = 0 \\ \lambda \geq 0 \end{array}$$

Karush-Kuhn-Tucker conditions

The Dual

$$\begin{array}{ll} \min_x & f(x) \\ \text{subject to} & g(x) \geq 0 \\ & h(x) = 0 \end{array}$$

The Primary problem

… is equivalent to the **dual problem**

$$\max_{\lambda \geq 0}(\min_x [f(x) - \lambda g(x) - \mu h(x)])$$

… that is, maximizing the Lagrangian over λ.

Sometimes it is easier to solve to dual … we will see this with SVMs.

A dual example

$$\begin{array}{ll} \min_x & f(x) = x^2 \\ \text{subject to} & x \geq 1 \end{array}$$

The Lagrangian function: $L(x, \lambda) = x^2 - \lambda(x - 1)$

The Dual function: $L^*(\lambda) = \min_x L(x, \lambda) = \min_x x^2 - \lambda(x - 1)$

Taking the derivative and setting to zero, the minimizer of the dual function is $x=\lambda/2$. Substituting, we get:

$$L^*(\lambda) = \lambda - \frac{1}{4}\lambda^2$$

The Dual problem: $\max_{\lambda \geq 0} \lambda - \frac{1}{4}\lambda^2$

Solution: λ=2, which is the Lagrange multiplier corresponding to x*=1

Some Optimization Methods

… over to you