- minimize $\frac{1}{2}(x_1^2 + 2x_2^2 + x_3^2)$

  subject to $x_1 + x_2 + x_3 = 5$

In this Chapter

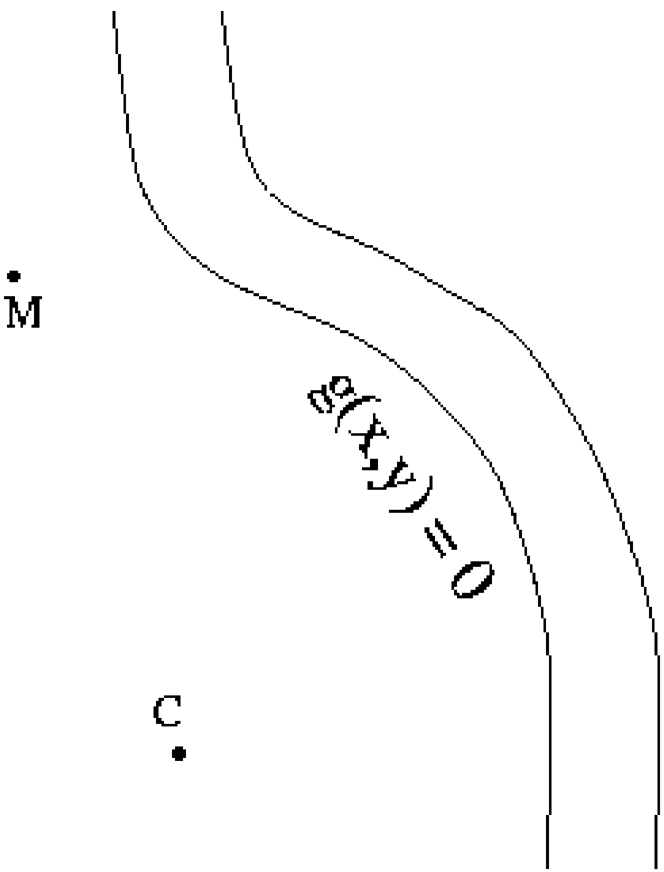
- We will study a new theory for constrained optimization
  - Local optimality condition
  - Easier to implement
  - Deeper insights to the geometrical structures
  - Explicit use of constraint functions
  - Convexity not required
  - Continuity and differentiability required

The methods I set forth require neither constructions nor geometric or mechanical considerations.

They require only algebraic operations subject to a systematic and uniform course.

Lagrange (1736 – 1813)

Can you help a lady in love?



A milkmaid is sent to get 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 **M**, she spots the cow at **C**. She wants to take the shortest path from where she is to the river and then to the cow.

An example problem

- minimize $f(x_1,x_2) = x_1 + x_2$

  subject to $h(x_1,x_2) = (x_1-1)^2 + x_2^2 - 1 = 0$
- What is x^*?
- What is $\nabla f(x^*)$?
- What is $\nabla h(x^*)$?
- What is the observation?

General Cases

- Consider an equality-constrained problem
  - Minimize $f(x)$
  - Subject to $h(x) = 0$
- Consider a local optimal point on the boundary
  - Geometric illustration of $\nabla f(x)$ and $\nabla h(x)$
  - **Demo**

A General Result

- If x^* is local minimum of f subject to $h(x) = 0$, then we can find a vector λ such that

$$\nabla f(x^*) + \sum_{i=1..m} \lambda_i \nabla h_i(x^*) = 0$$

But wait a minute…

- Is it always true?
- What are the assumptions?

Consider this problem:

- minimize $x_1 + x_2$

  s. t. $(x_1 - 1)^2 + x_2{}^2 - 1 = 0$

  $(x_1 - 2)^2 + x_2{}^2 - 4 = 0$
- How to visualize it?
  - What are the Lagrange multipliers?

What is wrong?

- When $\nabla h(x^*)$ are linearly dependent, i.e.
  - $\sum_{i=1..m} \lambda_i \nabla h_i(x^*) = 0$ for some λ_i

  The necessary condition may fail, since we cannot scale $\sum_{i=1..m} \lambda_i \nabla h_i(x^*)$ to match ∇f

- Regular point: a point x is regular if $\nabla h(x)$ are linearly independent

Lagrange Multiplier Theorem

- Let x^* be a local minimum and a regular point, then there exist unique scalars $\lambda^*_1, \ldots, \lambda^*_m$ such that

$$\nabla f(x^*) + \sum_{i=1..m} \lambda^*_i \nabla h_i(x^*) = 0$$

The Lagrange Condition

- A simple and elegant necessary condition
  - Counterpart of the first-order condition for unconstrained optimization
- Many different views to arrive at this

The Lagrangian Function

- Define the Lagrangian function

$$L(x, \lambda) = f(x) + \sum_{\mathrm{i=1..m}} \lambda_{\mathrm{i}} h_i(x)$$
$$= f(x) + \lambda' h(x)$$

- Then, if x^* is a local minimum and is regular, the Lagrange multiplier conditions are written as

$$\nabla_x L(x^*, \lambda^*) = 0, \ \nabla_\lambda L(x^*, \lambda^*) = 0$$

Let's try it out

- minimize $\frac{1}{2}(x_1^2 + 2x_2^2 + x_3^2)$
  subject to $x_1 + x_2 + x_3 = 5$
- Can we use the Lagrange Multiplier Theorem to solve the problem?

Let's try it out

- minimize $\quad -\frac{1}{2}(x_1^2 + 2x_2^2 + x_3^2)$

  subject to $\quad x_1 + x_2 + x_3 = 5$
- Can we use the Lagrange Multiplier Theorem to solve the problem?

A jewel box is to be constructed of material that costs \$1 per square inch for the bottom, \$2 per square inch for the sides, and \$5 per square inch for the top. If the total volume is to be 96 $in.^3$, what dimensions will minimize the total cost of construction?

An editor has been allotted $60,000 to spend on the development and promotion of a new book. It is estimated that if x thousand dollars is spent on development and y thousand on promotion, approximately $f(x, y) = 20x^{3/2}y$ copies of the book will be sold. How much money should the editor allocate to development and how much to promotion in order to maximize sales?

Second Order Sufficient Condition

- Let x^* and λ^* satisfy

$$\nabla_x L(x^*, \lambda^*) = 0, \ \nabla_\lambda L(x^*, \lambda^*) = 0,$$

$$y' \nabla^2_{xx} L(x^*, \lambda^*)\, y > 0, \ \forall y <> 0 \text{ with } \nabla h(x^*) y = 0.$$

Then x^* is a strict local minimum.

Let's try it out

- minimize $\quad -(x_1 x_2 + x_2 x_3 + x_3 x_1)$
  subject to $\quad x_1 + x_2 + x_3 = 3$
- Use the Lagrange multiplier theorem to answer the following:
  – What is x^* and λ^*?
  – Is x^* a local minimum or maximum?

(Sufficient condition for local minimum:

$y' \nabla^2_{xx} L(x^*, \lambda^*)\, y > 0,\ \forall y <> 0$ with $\nabla h(x^*) y = 0$)

Suppose in the editor Example, the editor is allotted $61,000 instead of $60,000 to spend on the development and promotion of the new book. Compute how the additional $1,000 will affect the maximum sales level.

Interpretation of λ^*

- Lagrange multipliers have an interesting interpretation
  - In economics, can be viewed as prices
- Consider the following problem:

  Minimize $f(x)$, subject to $a'x = b$
  - Suppose (x^*, λ^*) is a solution-multiplier pair
  - Suppose b is changed to $b+\Delta b$, the minimum x^* will change to $x^*+\Delta x$
  - What will be $f(x^*+\Delta x) - f(x^*)$ and how is the change related to λ^* ?

Interpretation of λ^* (cont'd)

- We have : $a'\Delta x = \Delta b$ and $\nabla f(x^*) = -\lambda^* a$,
- $\Delta \text{cost} = f(x^* + \Delta x) - f(x^*) = \nabla f(x^*)\Delta x + o(\Delta x) = -\lambda^* a'\Delta x + o(\Delta x)$
- Thus $\Delta \text{cost} = -\lambda^* \Delta b + o(\Delta x)$, so up to first order

$$\lambda^* = -\Delta \text{cost}/\Delta b$$

- For multiple constraints $a_i' x = b_i$, $i = 1, \ldots, m$, we have, in first order,

$$\Delta \text{cost} = -\sum_{i=1..m} \lambda^*_i \, \Delta b_i$$

- Conclusion: λ^* represents the rate of change of the optimal cost as the level of constraint changes