Intro to Optimization

Optimization and Pattern Recognition Fall, 2021

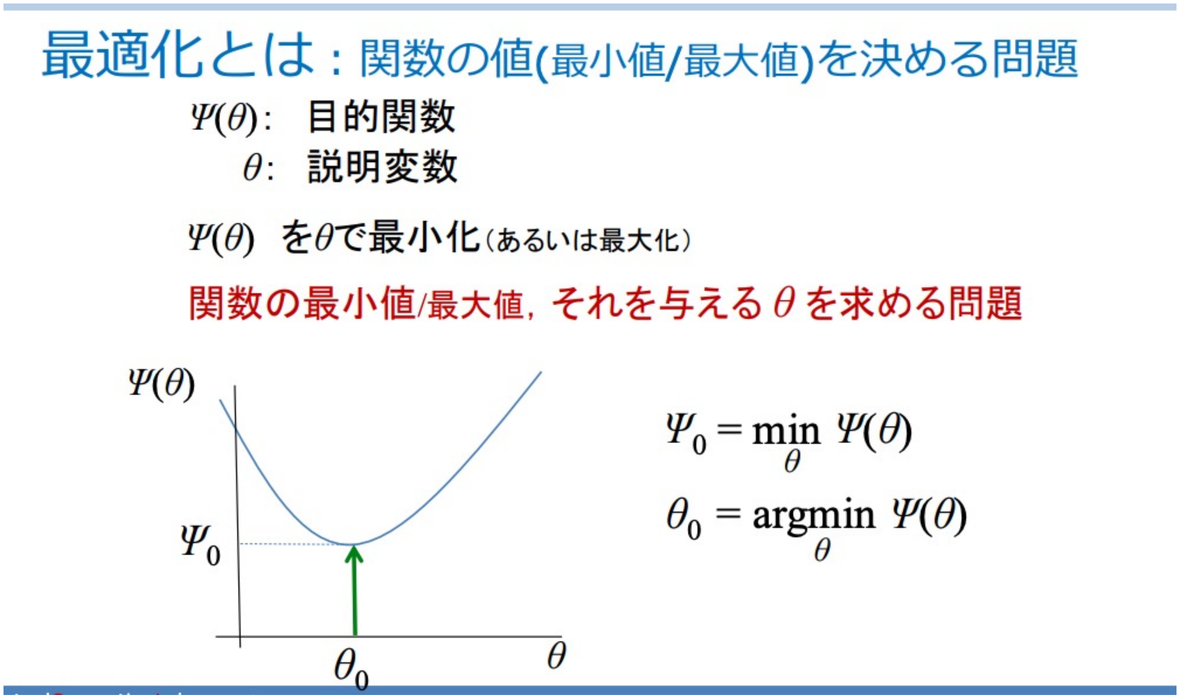
Jake Underland

2021-10-05

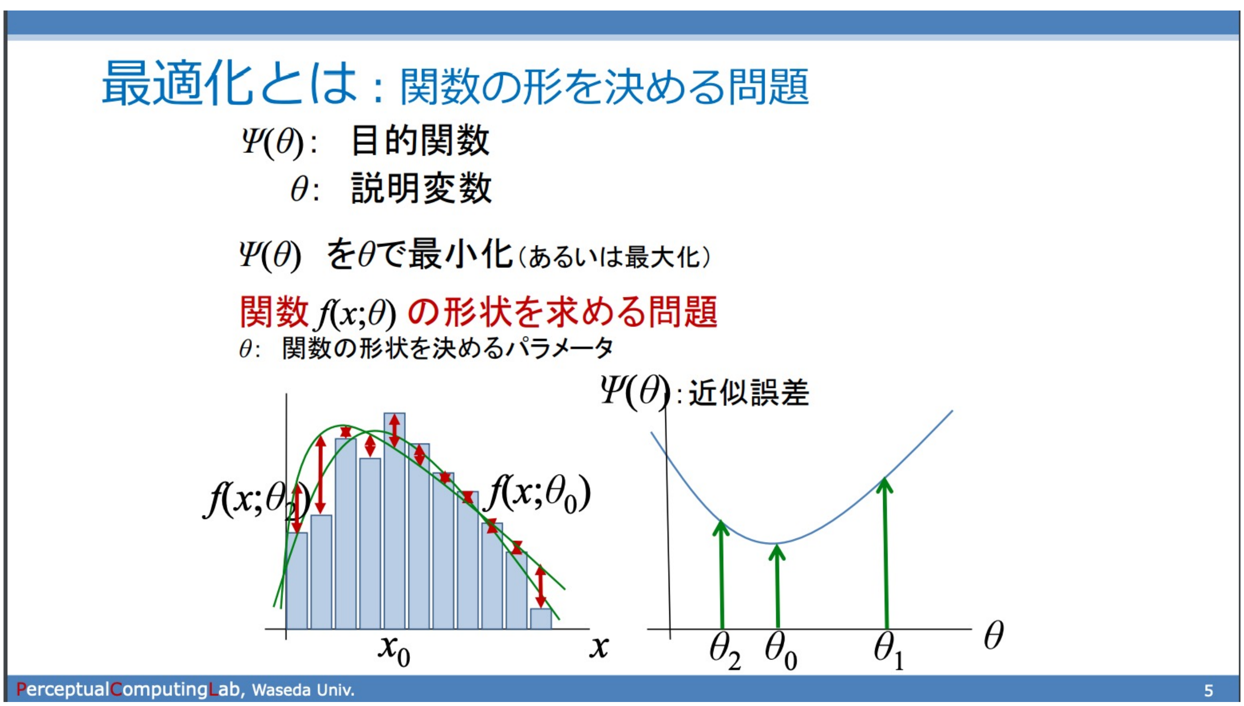
Contents

Types of Optimization

Finding Extremums

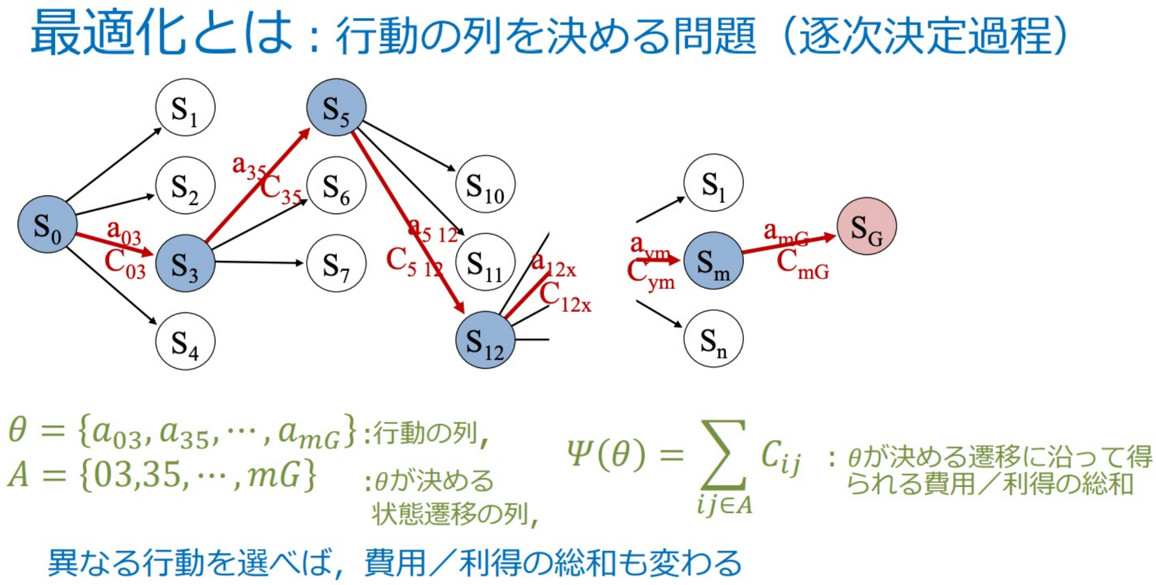


Determining Shape of Function



Find the value of variable θ_0 such that the error function $\Psi(\theta)$ is minimized. Note that this optimization problem can be translated into the first type (optimizing by finding extremums) through the use of an error function. Many approximation problems make use of this method.

Determining Optimal Sequence in Sequential Decision Process



θ represents a sequence of actions, A represents the sequence of states determined by θ, and $\Psi(\theta)$ is the cost of taking a certain θ, computed as $\sum_{ij \in A} C_{ij}$.

Formal Definition

$$\Psi(\theta) : \text{Objective (Target) Function}$$
$$\theta : \text{Explanatory Variable determining } \Psi$$

Optimization means to minimize (maximize) Ψ using θ.

Formal Characterization

Continuous Optimization

- **Nonlinear Programming Problem**:
  Includes nonlinearity in objective function or constraints.
  - *Quadratic Programming Problem*:
    Objective function is a quadratic function, constraints are linear
  - *Convex Programming Problem*:
    Objective function is a convex function, area of constraint is convex
- **Linear Programming Problem**:
  Objective function is linear, constraints are linear

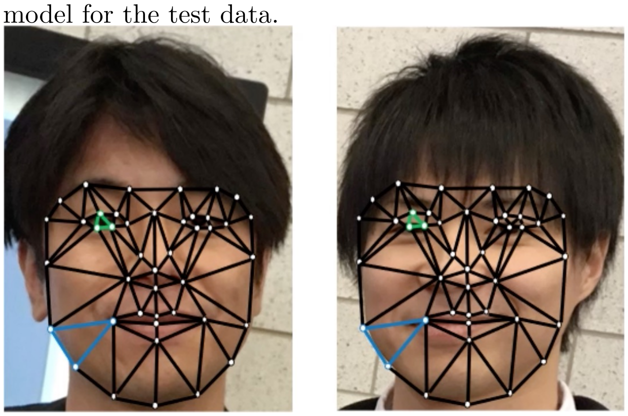
Discrete Optimization

- **Combinatorial Optimization Problem**
- **Integer Programming Problem**
  - *0-1 Integer Programming Problem:*
    Variables are either 0 or 1

Application Example: Facial Recognition

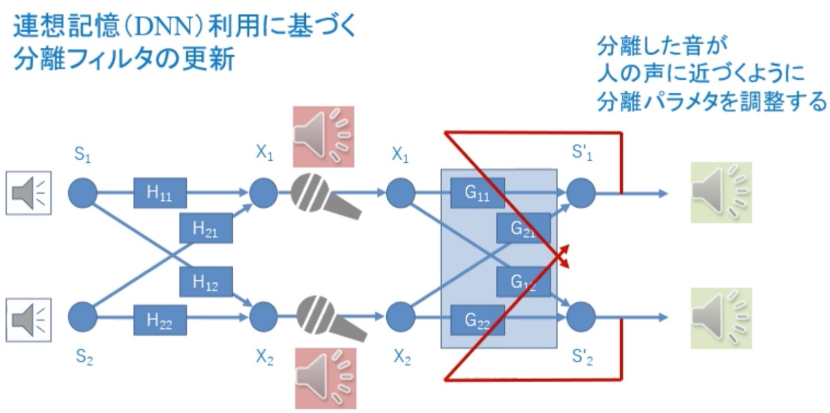
$$\theta : \text{Model for Facial Features}$$
$$\Psi(\theta) : \text{Pixel Error}$$

Begin by defining a model θ for facial features. The one below takes several prominent points of the face and connects them to form triangles. Then, we can define the error function $\Psi(\theta)$ to be the squared sum of pixel differences; that is, if a triangle on the training set captures black pixels but when applied to the testing set captures beige pixels, that is computed as an error. Then, the θ that minimizes $\Psi(\theta)$ provides us with a model for the test data.



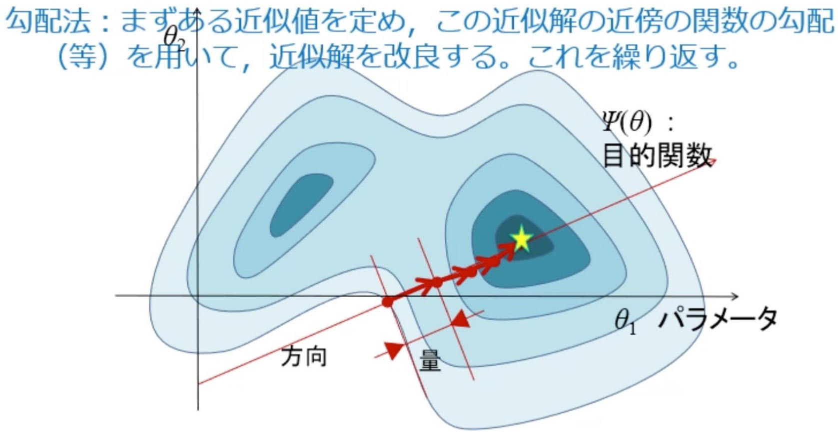
Application Example: Audio Source Separation

Two audio sources, two microphones, 4 paths (transformation functions) for the audio to travel. We can reverse this filter by feeding it back into the loop until each output audio sounds closer and closer to a normal human being.

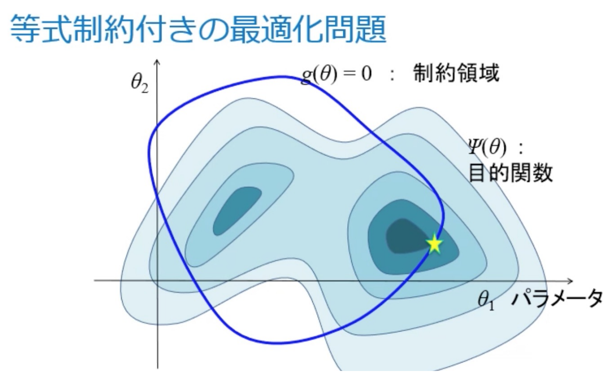


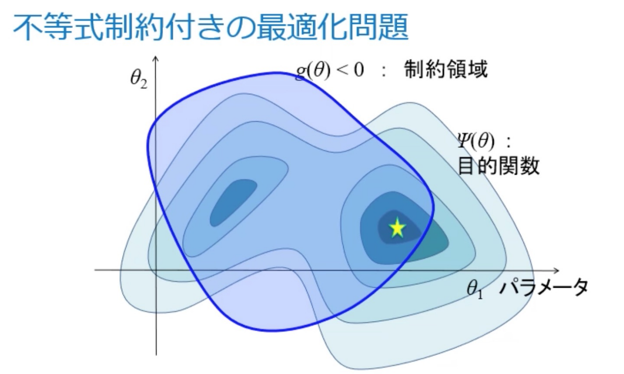
Solving Continous Optimization: Gradient Method

There is no analytic solution to continuous optimization, but one method is to approximate an initial value and study the gradient in the neighborhood of this value to arrive at the optimum.



We can add equality and inequality constraints of the parameters θ_1, θ_2 to the above:





Discrete Optimization Problems: