Constrained Optimization

Preview

constrained optimization을 배우기전에 사용했던 최적화로는

1. 미분해서0
2. Gradient descent

제약조건이 있을때는 어떻게 최적화?

Lagrange Mutliplier(라그랑주 승수법)

라그랑지 승수법(Lagrange multiplier) : 어떤 함수(F)가 주어진 제약식(h)을 만족시키면서, 그 함수가 갖는 최대값 혹은 최소값을 찾고자할 때 사용한다.

L(x,λ) = F(x) + λ*h(x) 으로 표기하며, **(x,λ)**변수들로 각각 편미분 한 식이 0이 되는 값으로 해를 구한다. 이러한 계산의 원리는, 사실 두 함수의 기울기가 같아지는 공통접선을 구하는 것이다.

- **부연 설명 : 라그랑지 승수법의 원리는 사실, 두가지 조건을 동시에 만족시키는 공통접선을 찾는 과정이다. 공통 접선이란 함수 F(x)와 제약식 h(x)을 미분해서 구한 접선의 기울기 벡터가 서로 평행한 점에서의 접선을 의미한다.**

즉, F(x)와 h(x)를 각각 미분하여 구한 기울기가 서로 평행하다는 것을 이용한다. 따라서 두 함수를 미분한 기울기에 앞뒤방향(+-)과 길이를 맞추기위한 미지수 λ를 곱하여, 이 두 접선이 서로 같다고 놓고 등식을 세운 것이다.

제약식(constraint)이 등호로 이루어진 경우→equality contraint problem

제약식이 부등호로 이루어진 경우→inequality constraint problem—>KKT condition이라는 라그랑지 승수법에 몇가지 조건을 추가하여 문제를 푼다.

제약식의 **λ가 0일경우 unconstraint problem이라 한다.**

Constrained Optimization

Formulation

f(x) is a function in n dimensional space,

that is , x=(x1,x2,x3,...xn)

-solve

$$\min_{\mathbf{x}} f(\mathbf{x})$$

$$\text{subjec to} \left\{ \begin{array}{l} g_j(\mathbf{x}) = 0 \;\; \text{for } j = 1,\Lambda\ ,m \\ h_j(\mathbf{x}) \le 0 \;\; \text{for } j = 1,\Lambda\ ,p \end{array} \right. \longrightarrow \text{constraints}$$

간단하게 하기 위해 g(x)와 h(x)는 convex 함수라 가정!

- **Solution of**

$$\min_{\mathbf{x}} f(\mathbf{x})$$
$$\text{subjec to} \begin{cases} g_j(\mathbf{x}) = 0 & \text{for } j = 1,\Lambda\ ,m \\ h_k(\mathbf{x}) \le 0 & \text{for } j = 1,\Lambda\ ,p \end{cases}$$

is equal to the solution of

KKT Multiplier

$$\min_{x} \max_{\boldsymbol{\lambda},\boldsymbol{\alpha}} F(\mathbf{x},\boldsymbol{\lambda},\boldsymbol{\alpha})$$
$$\text{subject to } \alpha_i \ge 0 \quad \text{for } i = 1,\Lambda\ ,p$$

where

$$F(\mathbf{x},\boldsymbol{\lambda},\boldsymbol{\alpha}) = f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x}) + \sum_{i=1}^{p} \alpha_i h_i(\mathbf{x})$$
$$\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \Lambda\ , \lambda_m)$$
$$\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \Lambda\ , \alpha_p)$$

Lagrange Multiplier

- **Solution of**

$$\min_{x} \max_{\boldsymbol{\lambda},\boldsymbol{\alpha}} F(\mathbf{x},\boldsymbol{\lambda},\boldsymbol{\alpha})$$
$$\text{subject to } \alpha_i \ge 0 \quad \text{for } i = 1,\Lambda\ ,p$$

$$\text{where } F(\mathbf{x},\boldsymbol{\lambda},\boldsymbol{\alpha}) = f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x}) + \sum_{i=1}^{p} \alpha_i h_i(\mathbf{x})$$

satisfies the following conditions

$$\frac{\partial}{\partial x_i} F(\mathbf{x},\boldsymbol{\lambda},\boldsymbol{\alpha}) = 0 \quad \text{for } i = 1,\Lambda\ ,n$$
$$g_j(\mathbf{x}) = 0 \quad \text{for } j = 1,\Lambda\ ,m$$
$$\alpha_j h_j(\mathbf{x}) = 0 \quad \text{for } j = 1,\Lambda\ ,p$$
$$h_j(\mathbf{x}) \le 0 \quad \text{for } j = 1,\Lambda\ ,p$$

Example

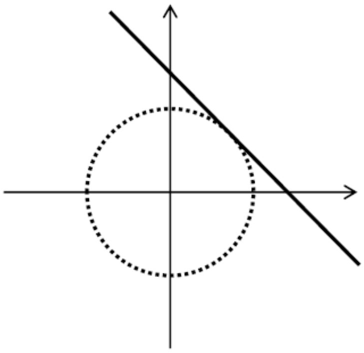
- **Example 1**

$$\min_{(x_1,x_2)} \left(x_1^2 + x_2^2\right)$$
$$\text{subjec to } x_1 + x_2 = 1$$

$\Rightarrow$

Find **x** which satisfies

$$\frac{\partial}{\partial x_1} F = 0$$
$$\frac{\partial}{\partial x_2} F = 0$$
$$x_1 + x_2 - 1 = 0$$
$$\text{where } F(x_1, x_2, \lambda_1) = x_1^2 + x_2^2 + \lambda_1(x_1 + x_2 - 1)$$



$$2x_1 + \lambda_1 = 0$$
$$2x_2 + \lambda_1 = 0$$
$$x_1 + x_2 - 1 = 0$$

$$x_1 = 0.5$$
$$x_2 = 0.5$$
$$\lambda_1 = -1$$

만약 , inequality constraint problem이 있을때

How to easily solve

1. Divides into sub-problems using KKT multipliers(알파)
2. Solve equations in each sub-problem
3. Check whether the solution satifies inequalities

- **Example 4**

$$\underset{(x_1,x_2)}{\arg\min}\left(x_1^2 + x_2^2\right) \quad \text{subjec to} \begin{cases} x_1 + x_2 = 1 \\ x_1 \ge 2 \\ x_2 \le 2 \end{cases} \Rightarrow$$

Find **x** which satisfies

$$\begin{aligned} 2x_1 + \lambda_1 - \alpha_1 &= 0 \\ 2x_2 + \lambda_1 + \alpha_2 &= 0 \\ x_1 + x_2 - 1 &= 0 \end{aligned} \qquad \begin{aligned} \alpha_1(-x_1 + 2) &= 0 \\ \alpha_2(x_2 - 2) &= 0 \\ -x_1 + 2 &\le 0 \\ x_2 - 2 &\le 0 \end{aligned}$$

Case $1: \alpha_1 = 0, \alpha_2 = 0$

$$\begin{aligned} 2x_1 + \lambda_1 &= 0 \\ 2x_2 + \lambda_1 &= 0 \\ x_1 + x_2 - 1 &= 0 \\ -x_1 + 2 &\le 0 \\ x_2 - 2 &\le 0 \end{aligned}$$

Case $2: \alpha_1 = 0, \alpha_2 \ne 0$

$$\begin{aligned} 2x_1 + \lambda_1 &= 0 \\ 2x_2 + \lambda_1 + \alpha_2 &= 0 \\ x_1 + x_2 - 1 &= 0 \\ x_2 - 2 &= 0 \\ -x_1 + 2 &\le 0 \end{aligned}$$

Case $3: \alpha_1 \ne 0, \alpha_2 = 0$

$$\begin{aligned} 2x_1 + \lambda_1 - \alpha_1 &= 0 \\ 2x_2 + \lambda_1 &= 0 \\ x_1 + x_2 - 1 &= 0 \\ -x_1 + 2 &= 0 \\ x_2 - 2 &\le 0 \end{aligned}$$

Case $4: \alpha_1 \ne 0, \alpha_2 \ne 0$

$$\begin{aligned} 2x_1 + \lambda_1 - \alpha_1 &= 0 \\ 2x_2 + \lambda_1 + \alpha_2 &= 0 \\ x_1 + x_2 - 1 &= 0 \\ -x_1 + 2 &= 0 \\ x_2 - 2 &= 0 \end{aligned}$$

- **Example 4**

Case $1: \alpha_1 = 0, \alpha_2 = 0$

$$\begin{aligned} 2x_1 + \lambda_1 &= 0 \\ 2x_2 + \lambda_1 &= 0 \\ x_1 + x_2 - 1 &= 0 \\ -x_1 + 2 &\le 0 \\ x_2 - 2 &\le 0 \end{aligned}$$

$$x_1 = 0.5, \quad x_2 = 0.5, \quad \lambda_1 = -1$$

Case $2: \alpha_1 = 0, \alpha_2 \ne 0$

$$\begin{aligned} 2x_1 + \lambda_1 &= 0 \\ 2x_2 + \lambda_1 + \alpha_2 &= 0 \\ x_1 + x_2 - 1 &= 0 \\ x_2 - 2 &= 0 \\ -x_1 + 2 &\le 0 \end{aligned}$$

$$x_1 = -1, \quad x_2 = 2, \quad \lambda_1 = 2, \quad \alpha_2 = 0$$

Case $3: \alpha_1 \ne 0, \alpha_2 = 0$

$$\begin{aligned} 2x_1 + \lambda_1 - \alpha_1 &= 0 \\ 2x_2 + \lambda_1 &= 0 \\ x_1 + x_2 - 1 &= 0 \\ -x_1 + 2 &= 0 \\ x_2 - 2 &\le 0 \end{aligned}$$

$$x_1 = 2, \quad x_2 = -1, \quad \lambda_1 = 2, \quad \alpha_1 = 6$$

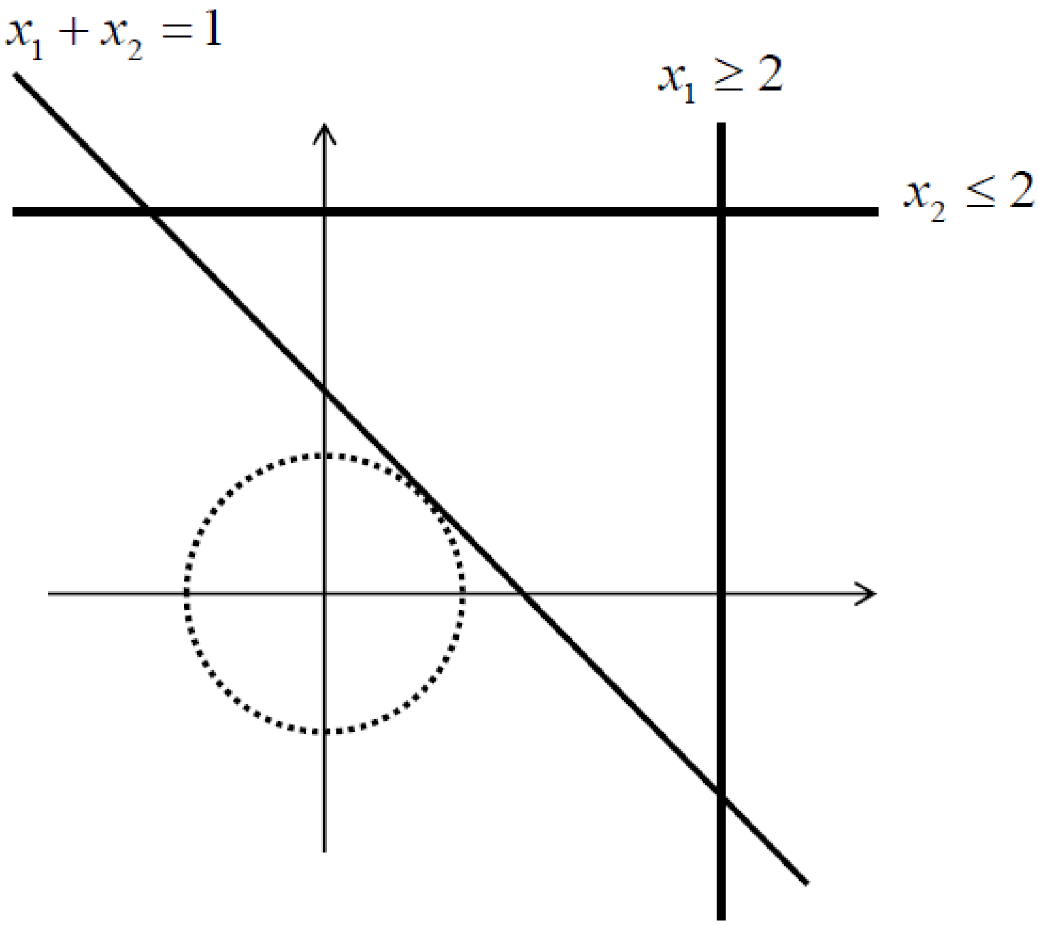
Case $4: \alpha_1 \ne 0, \alpha_2 \ne 0$

$$\begin{aligned} 2x_1 + \lambda_1 - \alpha_1 &= 0 \\ 2x_2 + \lambda_1 + \alpha_2 &= 0 \\ x_1 + x_2 - 1 &= 0 \\ -x_1 + 2 &= 0 \\ x_2 - 2 &= 0 \end{aligned}$$

No solution!

case1 만족 X, case2 만족 X, case3 만족함

Example4를 그림으로 표현하면 다음과 같다.



Lagrange Multiplier Methods for Constrained Optimization

-Good When

f(x) is 2nd order

g(x)and h(x) are linear

-If there are P inequality conditions

- You have to slove 2^p sets of simultaneous equations in the worst case
- Need combinatorial search

Dual Form

– Solution of

$$\min_{\mathbf{x}} \max_{\boldsymbol{\lambda},\boldsymbol{\alpha}} F(\mathbf{x},\boldsymbol{\lambda},\boldsymbol{\alpha})$$
$$\text{subject to } \alpha_i \geq 0 \quad \text{for } i = 1,\Lambda\ , p$$

is equal to the solution of

$$\max_{\boldsymbol{\lambda},\boldsymbol{\alpha}} \min_{\mathbf{x}} F(\mathbf{x},\boldsymbol{\lambda},\boldsymbol{\alpha})$$
$$\text{subject to } \alpha_i \geq 0 \quad \text{for } i = 1,\Lambda\ , p$$

We will use this form for SVM