

# AI 알고리즘

## 비선형계획법

# 제약 조건이 없는 비선형 계획 모형

비선형 계획 모형

## 비선형계획모형

### ❖ 볼록성과 오목성

#### ■ 정리1

- 만약 함수  $f(x)$ 가  $S$ 에서 오목하다면 이 비선형계획문제에 대한 어떤 국부 극댓값은 이 문제의 최적해가 됨
  - 가장 높은 봉우리를 찾는 것

#### ■ 정리2

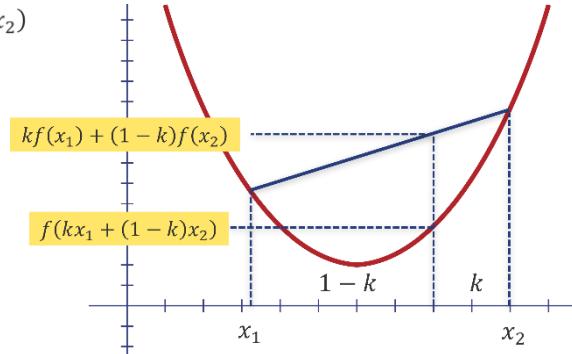
- 만약 함수  $f(x)$ 가  $S$ 에서 볼록하다면 이 비선형계획문제에 대한 어떤 국부 극솟값은 이 문제의 최적해가 됨
  - 가장 낮은 골짜기가 극솟값이며 최적해

#### ■ 볼록함수

- 함수  $f(x)$ 가  $x$ 의 정의역 내에 있는 임의의  $x_1, x_2$ 에 대해서 다음 조건을 만족하면 볼록함수라고 함

$$f(kx_1 + (1 - k)x_2) \leq kf(x_1) + (1 - k)f(x_2)$$

단,  $k$ 는  $0 < k < 1$ 인 실수

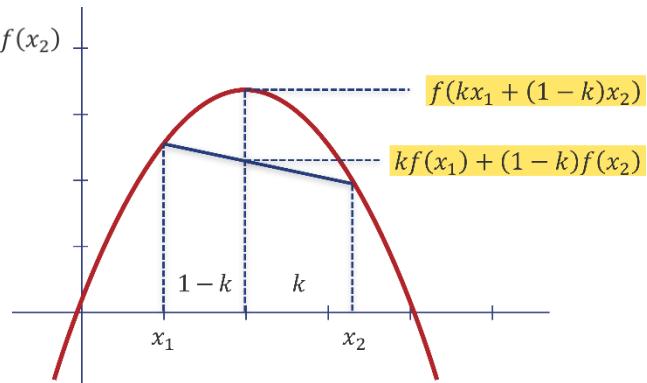


#### ■ 오목함수

- 함수  $f(x)$ 가  $x$ 의 있는 임의의  $x_1, x_2$ 에 대해서 다음 조건을 만족하면 오목함수라고 함

$$f(kx_1 + (1 - k)x_2) \geq kf(x_1) + (1 - k)f(x_2)$$

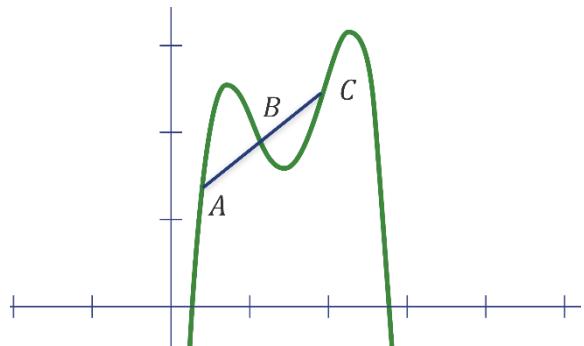
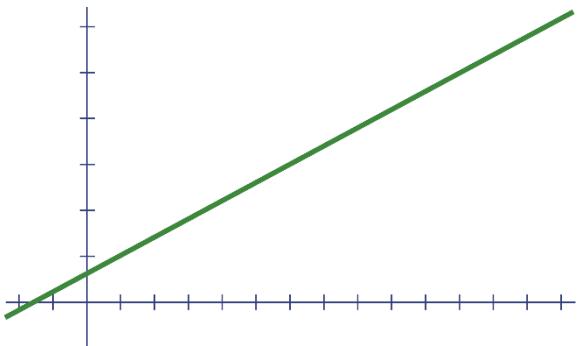
단,  $k$ 는  $0 < k < 1$ 인 실수



## 비선형계획모형

### ❖ 볼록성과 오목성

- 볼록함수&오목함수
- 볼록함수아님&오목함수아님



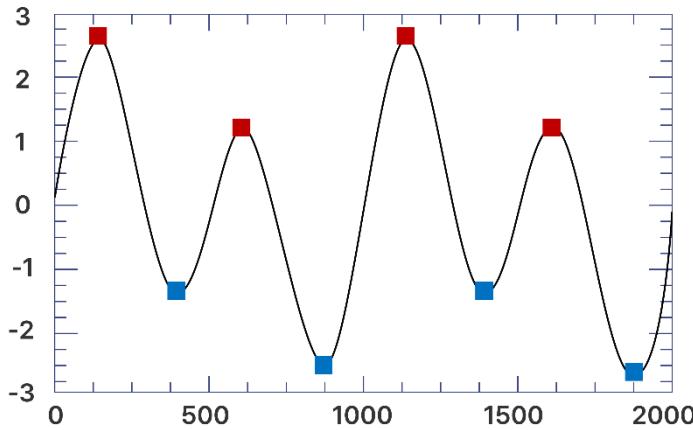
### ❖ 국부극값(Local Extremum)

- 어떤 최대화의 비선형계획문제
  - $|x_i - x'_i| < \epsilon$ (매우 작은 값)를 가지고 있는 어떤 실행 가능점  $x' = (x'_1, x'_2, \dots, x'_n)$ 이  $f(x) \geq f(x')$ 를 보인다면  
→  $x = (x_1, x_2, \dots, x_n)$ 을 국부극댓값이라고 함
- $x$ 와 가까운 실행 가능한  $x'$ 에 대해  $f(x) \geq f(x')$ 이라면 점  $x$ 는 국부극댓값
- 최소화문제에서는  $x$ 와 가까운 실행 가능한  $x'$ 에 대해  $f(x) \leq f(x')$ 이라면 점  $x$ 는 국부극솟값
- 주변에 있는 점보다 작은 점은 그로컬에서 극솟값이며, 주변 점보다 큰 점은 극댓값
- 최대화 선형계획문제에서는 어떤 국부극댓값이 선형계획문제의 최적해이지만 비선형계획문제의 경우에는 그렇지 않음

## 비선형계획모형

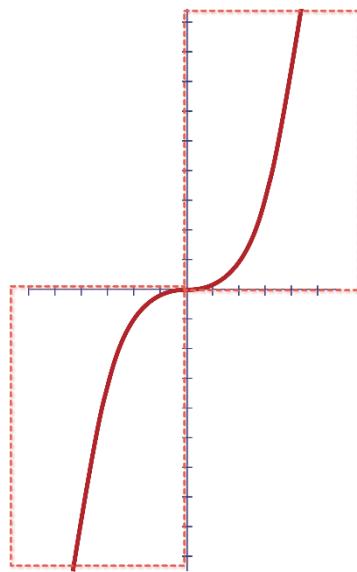
### ❖ 비선형목적함수를 가지는 최적화문제의 극값

- 가장 높은 봉우리가 무엇인가?



### ❖ 2차 도함수를 이용하여 볼록(오목)함수 결정방법

- $x_1 \leq x \leq x_2$  일 때, 함수  $f(x)$ 에 대해
  - 물리적 모형
  - $f''(x) \geq 0$  이면 볼록함수 ( $f''(x)$  가  $x_1, x_2$  구간에서 연속적)
  - $f''(x) \leq 0$  이면 오목함수 ( $f''(x)$  가  $x_1, x_2$  구간에서 연속적)
- $f(x) = x^3$ 
  - $f'(x) = 3x^2, f''(x) = 6x$  ( $x$  범위에 따라 다름)



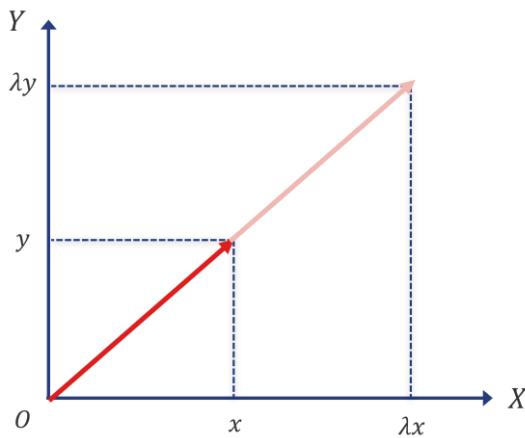
# 제약 조건이 없는 비선형 계획 모형

Hessian 행렬식

## Hessian 행렬식

### ❖ Hessian 행렬식

- 기본적인 지식을 가지고 어떻게 푸는가?
- 고윳값과 고유벡터
  - 특정한 벡터와 행렬은 선형 변환을 취해주었을 때, 크기만 바뀌고 방향은 바뀌지 않을 수도 있음



- 그렇다면, 그 크기는 얼마만큼 변했나?
- $A\mathbf{x} = \lambda\mathbf{x}$ 
  - $(A - \lambda I)\mathbf{x} = 0$
  - $\det(A - \lambda I) = 0$
- 행렬식(Determinant)
  - 정사각 행렬에서 만정의 되는 값으로, 행렬의 가역성을 판별해 줌
  - 표기법

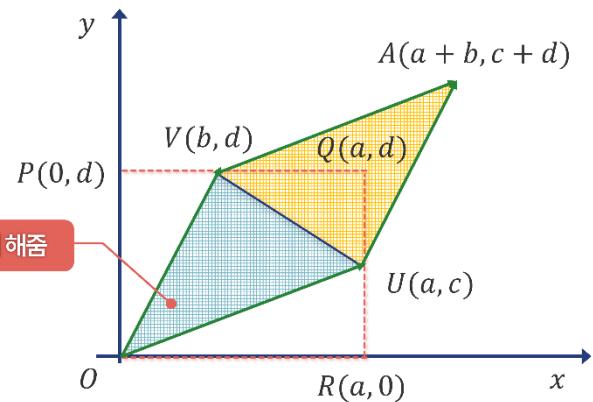
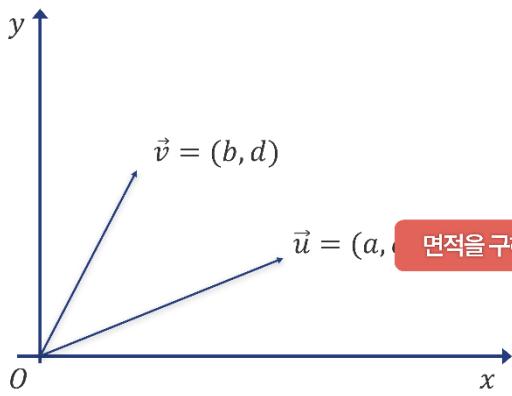
$$\begin{aligned}|A| &= \begin{vmatrix} a & b \\ c & d \end{vmatrix} = \left| \begin{pmatrix} a & b \\ c & d \end{pmatrix} \right| \\ &= \det A = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc\end{aligned}$$

- 다른 공간으로 이동했지만, 같은 성질을 가지는 벡터
- 선형변환 후의 두 기저 벡터가 이루고 있는 평행사변형의 넓이
  - 행렬은 일종의 선형변환  $(1, 0), (0, 1) \rightarrow (a, c), (b, d)$

## Hessian 행렬식

### ❖ Hessian 행렬식

- 선형변환 후의 두 기저 벡터가 이루고 있는 평행사변형의 넓이
  - 면적의 변화가 바로 행렬식(Determinant)

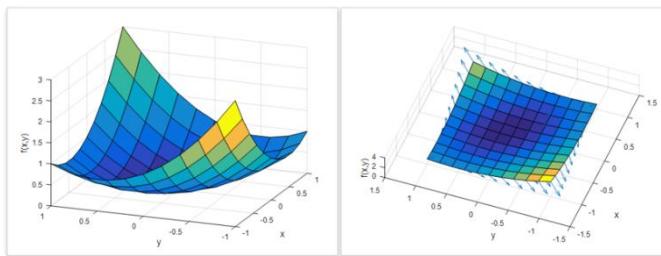


- 편미분(Differentiation)
  - 함수에서 미분의 개념은 기울기를 의미
  - 여러 개의 변수를 가짐
  - 다변수 함수에서 임의의 점 하나를 찍어 보면 그 점에서 기울기 직선은 하나로 결정되지 않음
  - 수직인 두 방향으로만 기울기를 구하기 위하여 필요
    - $f(x, y)$ 의 x축으로의 기울기와 y축으로의 기울기를 각각 구할 수 있음
    - y 또는 x를 상수로 놓고 미분하여 x방향만으로의 기울기와 y방향만으로의 기울기를 구함
  - 함수  $f$ 의 x방향으로의 편미분
    - $f_x$  혹은  $\frac{\partial f}{\partial x}$
  - 함수  $f$ 의 y방향으로의 편미분
    - $f_y$  혹은  $\frac{\partial f}{\partial y}$

## Hessian 행렬식

### ❖ Hessian 행렬식

- 기울기(Gradient)
  - x방향으로의 편미분 값과 y방향으로의 편미분 값을 원소로 하는 벡터로 출력
  - 즉, 기울기를 각각 벡터 형태로 만들어 놓은 것
  - 스칼라 함수로부터 벡터장을 형성하는 연산자
  - 해당 포인트에서  $f(x,y)$  평면의 기울기가 가장 가파른 곳으로의 방향을 의미



- Hessian 행렬
  - 어떤 함수의 2계도 함수들을 이용하여 행렬을 만든 것
  - 실함수  $f(x_1, x_2, \dots, x_n)$ 의 Hessian 행렬
  - 특정 고유벡터에 대해 고윳값의 크기가 클수록 해당 방향으로 더 볼록한 것
  - 고윳값이 모두 양수라면 함수는 아래로 볼록하며, 이것이 임계점이라면 극솟값
  - 모두 음수라면 함수는 위로 볼록할 것이며, 이것이 임계점이라면 극댓값
  - 주소 행렬식(Principal Minor)
    - Hessian 행렬의  $n-i$  열에 대응하는  $n-i$  행을 제거함으로써 얻어진  $i \times i$  행렬의 행렬식
    - $M_1 \geq 0, M_2 \geq 0, \dots, M_n \geq 0$  이면 볼록
    - $M_1 \leq 0, M_2 \geq 0, \dots, M_n(-1)^n \geq 0$  이면 오목
- Hessian 행렬 활용 이유
  - 특정 위치가 위로 볼록한지, 아래로 볼록한지, 안장점인지를 판단할 수 있음

$$H(f) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \ddots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

# 제약 조건이 없는 비선형계획모형

제약 조건이 없는 비선형계획모형

## 제약조건이 없는 비선형계획모형

### ❖ 일변수함수의 최적화

- 제약식이 없고, 두 번의 미분만 있으면 됨
- 점  $x_0$ 가 일변수함수  $f(x)$ 의 극점이 되는 필요조건
  - $\frac{(\partial f(x_0))}{\partial f(x)} = 0$
- 점  $x_0$ 에서 함수의 기울기는 0
- 그런 점을 정점(Stationary Point)이라고 함
- 정점은 극대점, 극소점, 변곡점이 됨
  - $\frac{\partial^2 f(x_0)}{\partial x^2} > 0$ 이면  $x_0$ 에서 함수  $f(x)$ 가 국부극솟값을 갖는 충분조건
  - $\frac{\partial^2 f(x_0)}{\partial x^2} < 0$ 이면  $x_0$ 에서 함수  $f(x)$ 가 국부극댓값을 갖는 충분조건
- $\frac{\partial^2 f(x_0)}{\partial x^2} = 0$ 이면 고차도함수를 구하여  $\frac{\partial^{n-1} f(x_0)}{\partial x^{n-1}} = 0$ 이고,  $\frac{\partial^n f(x_0)}{\partial x^n} \neq 0$ 일 때
  - $n$ 이 홀수이면 변곡점
  - $n$ 이 짝수이고,  $\frac{\partial^n f(x_0)}{\partial x^n} > 0$ 이면  
함수  $f(x)$ 는 국부극솟값
  - $n$ 이 짝수이고,  $\frac{\partial^n f(x_0)}{\partial x^n} < 0$ 이면  
함수  $f(x)$ 는 국부극댓값

## 제약조건이 없는 비선형계획모형

### ❖ 일변수함수의 최적화 예제

- $f(x) = x^2 - 4x + 8$ 
  - $\frac{\partial f(x_0)}{\partial f(x)} = 2x - 4 = 0, x = 2$
  - $\frac{\partial^2 f(x_0)}{\partial x^2} = 2 > 0$
  - $f(x)$ 는 볼록함수이고,  $x = 2$ 는 국부극소값
  
- $f(x) = x^3 - 3x$ 
  - $\frac{\partial f(x_0)}{\partial f(x)} = 3x^2 - 3 = 0, x = \pm 1$
  - $\frac{\partial^2 f(x_0)}{\partial x^2} = 6x,$ 
    - $x > 0$ 이면 국부극소값 ( $x=1$ )
    - $x < 0$ 이면 국부극대값 ( $x=-1$ )
    - $x=0$ 이면  $\frac{\partial^2 f(x_0)}{\partial x^2} = 0$ 이므로 고차도함수를 구함
    - $\frac{\partial^3 f(x_0)}{\partial x^3} = 6 > 0$ 이고,  $n$ 은 홀수이므로  $x=0$ 은 변곡점

### ❖ 이변수함수의 최적화

- 이변수함수  $f(x_1, x_2)$ 에서 극값을 갖게 되는 필요조건

$$\frac{\partial f(x_1, x_2)}{\partial f(x_1)} = 0, \quad \frac{\partial f(x_1, x_2)}{\partial f(x_2)} = 0$$

- 위 두식을 만족하는  $x^* = (x_1^*, x_2^*)$ 는 정점
- 정점에서 Hessian 행렬의 주소행렬식  $M_1, M_2$ 를 구함
  - $M_1 > 0, M_2 > 0$ 이면  $(x_1^*, x_2^*)$ 는 국부극솟값
  - $M_1 < 0, M_2 > 0$ 이면  $(x_1^*, x_2^*)$ 는 국부극댓값
  - $M_2 < 0$ 이면 변곡점
  - $M_2 = 0$ 이면 추가적인 분석이 필요

## 제약조건이 없는 비선형계획모형

### ❖ 이변수함수의 최적화 예제

$$f(x_1, x_2) = x_1^3 + x_1x_2 + x_2^2$$

$$\frac{\partial f(x_1, x_2)}{\partial f(x_1)} = 3x_1^2 + x_2 = 0, \frac{\partial f(x_1, x_2)}{\partial f(x_2)} = x_1 + 2x_2 = 0 \quad M_1 = \frac{\partial f^2}{\partial x_1^2}$$

$$x^* = (x_1^*, x_2^*) = (0,0), (1/6, -1/12)$$

(1/6, -1/12)는 국부극솟값

$$M_1 = (3x_1^2 + x_2)' = 6x_1 = 1 > 0$$

$$M_2 = 12x_1 - 1 = 1 > 0$$

$$M_2 = \begin{vmatrix} \frac{\partial f^2}{\partial x_1^2} & \frac{\partial f^2}{\partial x_1 \partial x_2} \\ \frac{\partial f^2}{\partial x_2 \partial x_1} & \frac{\partial f^2}{\partial x_2^2} \end{vmatrix}$$

(0,0)는 변곡점

$$M_2 = -1 < 0$$

$$M_2 = \begin{vmatrix} 6x_1 & 1 \\ 1 & 2 \end{vmatrix}$$

### ❖ 다변수함수의 최적화

- 함수  $f(x_1, x_2, \dots, x_n)$  가  $x^* = (x_1^*, x_2^*, \dots, x_n^*)$  에서 극값을 갖게 되는 필요조건
  - 기울기라는 의미가 편미분

$$\frac{\partial f}{\partial x_1} = 0, \quad \frac{\partial f}{\partial x_2} = 0, \dots, \frac{\partial f}{\partial x_n} = 0$$

- $x^* = (x_1^*, x_2^*, \dots, x_n^*)$  가 극값을 갖게 되는 충분조건
  - $M_1 > 0, M_2 > 0, \dots, M_n \geq 0$  이면 국부극솟값
  - $M_1 < 0, M_2 > 0, \dots, M_n (-1)^n > 0$  이면 국부극댓값
  - 위와 다른 형태가 나타날 경우 복잡한 분석이 필요

### ❖ 일반적인 비선형함수에 적용

- 최적해에서는 비선형함수의 1차 미분함수 기울기가 0이 되어야 함
- 이해의 최소값, 최대값 여부를 판단하기 위해서는 2차 미분함수의 기울기를 파악하여야 함

## 제약조건이 없는 비선형계획모형

## 제약조건이 없는 비선형계획모형