04.05 변분법

- 범함수 : 입력이 함수인 것

  ex) F[f(x)]

05.01 최적화 기초

1) 최적화 문제

- 최적화 문제 : 함수의 값을 최대화 혹은 최소화하는 변수 x의 값을 찾는 것
- 최적화의 대상이 되는 함수 : 목적함수 (loss, cost, error function)

2) 그리드 서치와 수치적 최적화

- 그리드 서치 : 가장 간단한 방법. 계산량이 많다는 단점.

가능한 x값을 모두모두 넣어 최적점을 찾는 방법

- 수치적 최적화 : trial and error를 통해 최적점을 찾는 방법.

가능한 한 적은 횟수로 탐색하는 방법

[필요한 알고리즘]

1) 최적점인지 판단 : 기울기 필요조건 (1차 미분 : 0, 2차 미분 > 0 이면 최소값)

*행렬의 경우 = 야코비안(1차 미분), 헤시안(2차 미분) 으로 판단

$J = 0$ H가 양의정부호

2) 최적점이 아니라면, 다음 번에 시도할 위치를 찾는 방법

(최대경사법, 뉴턴방법, 준뉴턴방법(BFGS))

05.01 최적화 기초 | 3) 최대경사법(steepest gradient descent)

- 현재 위치에서 기울기 값 만을 이용하여 다음번 위치를 결정

- (5.1.11)
  - *기울기의 부호가 (-)라는 점 주의. 기울기의 (-)방향으로 진행.
  - *하산하고 싶을 때, 경사가 오르막(+)이라면, 반대방향으로 내려가야 함
  - ’기울기’가 음수면, x를 +방향으로 전진시킴
  - '뮤'는 옮기는 거리의 비례상수로, step size라고 한다. 내려가는 걸음폭.

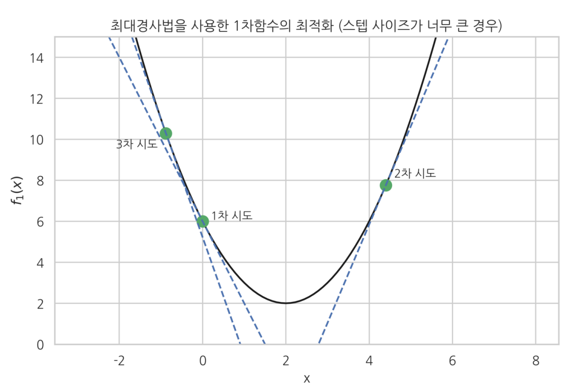
최대경사법(Steepest Gradient Descent)방법은 단순히 현재 위치 x_k 에서의 기울기 값 $g(x_k)$ 만을 이용하여 다음번 위치 x_{k+1} 를 결정하는 방법이다.

$$x_{k+1} = x_k - \mu \nabla f(x_k) = x_k - \mu g(x_k) \tag{5.1.11}$$

만약 현재 위치 x_k 에서 기울기가 음수면 즉 곡면이 아래로 향하면 $g(x_k) < 0$이므로 앞으로 진행하고 현재 위치 x_k 에서 기울기가 양수면 $g(x_k) > 0$이므로 뒤로 진행하게 되어 점점 낮은 위치로 옮겨간다. 이때 위치를 옮기는 거리를 결정하는 비례상수 μ를 **스텝 사이즈(step size)**라고 한다.

x_k 가 일단 최적 점에 도달했을 때는 $g(x_k) = 0$이 되므로 더 이상 위치를 옮기지 않는다.

- step size를 너무 크게 한다면, 생기는 문제는? (05.01 9page)



- 곡면의 모양이 '계곡'처럼 생긴 경우, 어떠한 현상이 발생하고, 문제점은 무엇인가?

연산↑ 진동

- 위의 현상을 없애기 위한 방법으로 어떠한 방법이 사용되는 가?

H를 활용한
뉴턴방법

4) 2차 도함수를 사용한 뉴턴 방법

- (5.1.13) *경사와 보폭(기울기와 stepsize)를 헤시안행렬-1 로 대체함
- 뉴턴 방법으로 보완한 2가지 (보폭과 경사)

$$x_{n+1} = x_n - [Hf(x_n)]^{-1} \nabla f(x_n)$$

$x_{n+1} = x_n - \frac{f'(x_n)}{f''(x_n)}$

$x_{n+1} = x_n - H^{-1} \nabla f(x_n)$

*최적의 stepsize가 $\frac{1}{f''(x_n)}$ or H^{-1}

5) 준 뉴턴방법

- 헤시안 행렬을 직접 구하기 어려울 때, 사용
- 그레디언트를 바로 적용하지 않고, 그레디언트에 2차 도함수 관련된 무언가를 넣어서 계산

6) 전역 최적화 문제

ML에선, 대부분 convex function으로, 전역 최적화 가능.

- 복수의 local minimum을 갖고 있을 때, 초기값이나 step size를 잘못 주게 되면, global minimum을 찾지 못하고 종료되버리는 문제

- 딥러닝에선 해당되지 않지만, 머신러닝에선 전역 최적화 문제가 아래의 조건을 만족하면 '컨벡스 문제'로 귀결된다.

- 그리고, 컨벡스 문제는 항상 global minimum을 갖고, 이를 거의 찾을 수 있다. (컨벡스 문제가 아닌 딥러닝에서 골치가 아플 수도 있음)

  * 컨벡스 문제일 조건 :

2차 도함수 >= 0,   H가 양의 준정부호 (다변수 함수일 경우)

05.02 제한조건이 있는 최적화 문제

1) 제한조건 : 연립방정식 or 연립부등식

- 그냥 x*를 찾는 게 아니라, 제한조건을 만족하는 x*를 찾아라!
  - 연립방정식 제한조건 : 라그랑지 승수법 사용해 최적화 해결
  - 연립부등식 제한조건 : KKT 조건을 만족하도록 해야함

2) 등식 제한조건 (05.02 1page)

현실의 최적화 문제에서는 여러가지 **제한조건이 있는 최적화(constrained optimization)** 문제가 많다. 가장 간단한 경우는 다음과 같이 연립방정식 제한조건이 있는 경우다. **등식(equality) 제한조건**이라고도 한다.

$$x^* = \arg\min_x f(x) \tag{5.2.1}$$

$$x \in \mathbf{R}^N \tag{5.2.2}$$

$$g_j(x) = 0 \quad (j = 1, \ldots, M) \tag{5.2.3}$$

첫 번째 식만 보면 단순히 목적함수 $f(x)$를 가장 작게 하는 N차원 벡터 x값을 찾는 문제다. 하지만 마지막 식에 있는 M개의 등식 제한 조건이 있으면 M개 연립 방정식

$$\begin{aligned} g_1(x) &= 0 \\ g_2(x) &= 0 \\ &\vdots \\ g_M(x) &= 0 \end{aligned} \tag{5.2.4}$$

를 동시에 모두 만족시키면서 목적함수 $f(x)$를 가장 작게 하는 x값을 찾아야 한다.

3) 라그랑지 승수법

이렇게 등식 제한조건이 있는 최적화 문제는 **라그랑주 승수법(Lagrange multiplier)**을 사용하여 최적화할 수 있다.

라그랑주 승수 방법에서는 목적함수를 원래의 목적함수 $f(x)$를 사용하지 않는다. 대신 제한조건 등식에 λ라는 새로운 변수를 곱해서 더한 함수

$$\begin{aligned} h(x, \lambda) &= h(x_1, x_2, \ldots, x_N, \lambda_1, \ldots, \lambda_M) \\ &= f(x) + \sum_{j=1}^{M} \lambda_j g_j(x) \end{aligned} \tag{5.2.7}$$

를 목적함수로 간주하여 최적화한다. 이때 제한조건 등식 하나마다 새로운 λ_i를 추가해주어야 한다. 따라서 만약 제한조건이 M개이면 $\lambda_1, \cdots, \lambda_M$개의 변수가 새로 생긴 것과 같다. 이렇게 확장된 목적함수 h는 입력변수가 더 늘어났기 때문에 그레디언트 벡터를 영벡터로 만드는 최적화 필요 조건이 다음처럼 $N + M$개가 된다.

$$\begin{aligned} \frac{\partial h}{\partial x_1} &= \frac{\partial f}{\partial x_1} + \sum_{j=1}^{M} \lambda_j \frac{\partial g_j}{\partial x_1} = 0 \\ \frac{\partial h}{\partial x_2} &= \frac{\partial f}{\partial x_2} + \sum_{j=1}^{M} \lambda_j \frac{\partial g_j}{\partial x_2} = 0 \\ &\vdots \\ \frac{\partial h}{\partial x_N} &= \frac{\partial f}{\partial x_N} + \sum_{j=1}^{M} \lambda_j \frac{\partial g_j}{\partial x_N} = 0 \\ \frac{\partial h}{\partial \lambda_1} &= g_1 = 0 \\ &\vdots \\ \frac{\partial h}{\partial \lambda_M} &= g_M = 0 \end{aligned} \tag{5.2.8}$$

05.02 제한조건이 있는 최적화 문제

05.02 제한조건이 있는 최적화 문제

3) 라그랑지 승수법 (라그랑지 승수 = 람다)

0) 제한식 = 0 꼴로 다듬기 (gi(x) = 0)

1) 등식 제한조건에 람다를 붙여 새로운 목적함수 생성

2) 기존 변수(x들)와 람다변수들 모두의 그레디언트벡터 = 0 으로 하는 xi*, 람다i* 를 찾는다. (최적해)

*람다i = 0 이라면, 해당 제한등식(gi(x))는 무의미한 제한식

* 라그랑지 승수(람다)의 의미

: 제한조건이 의미가 있으려면 람다는 0이 아니다.

(제한조건 유무에 관계없이 해가 같다면, 람다=0, 제한조건 없이 그냥 최적화 진행과 같음)

(람다=0 일때만 등식이 있을 때와 없을 때의 최적해가 같아짐)

4) 부등식 제한조건

- 제한조건인 부등식은 g(x)<= 0 으로 부등호 방향이 되도록 해야함
- 라그랑지 승수법을 적용해 새로운 목적함수 만들어 최적화 실시.

단, KKT 조건을 만족해야 최적해 찾을 수 있음

- KKT 조건 (05.02 7page)

(1) 모든 독립 변수 $x_1, x_2, \ldots, x_N$에 대한 미분값이 0이다.

$$\frac{\partial h(x, \lambda)}{\partial x_i} = 0 \quad (5.2.31)$$

$\frac{\partial h(x,\lambda)}{\partial \lambda_i}$ 는 따지지 않음

(2) 모든 라그랑주 승수 $\lambda_1, \ldots, \lambda_M$과 제한조건 부등식($\lambda$에 대한 미분값)의 곱이 0이다. ⟺ 실상은 제한조건X 아 등식제한조건

$$\lambda_j \cdot \frac{\partial h(x, \lambda)}{\partial \lambda_j} = \lambda_j \cdot g_j = 0 \quad (5.2.32)$$

$\lambda = 0$ 아 $\frac{\partial h}{\partial \lambda_i} = 0$

$\lambda = 0$이라면, 제한조건 없이 최적화 푸는것과 같음. $\frac{\partial h}{\partial \lambda_i} = 0$ 이면 등식제한조건이 된다.

(3) 라그랑주 승수는 음수가 아니어야 한다.

$$\lambda_j \geq 0 \quad (5.2.33)$$

1) 다변수에 대한 미분 값 = 0

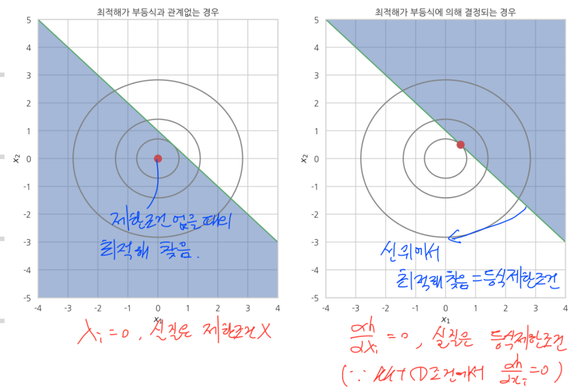
<=> x로 미분한 값 = 0

2) 라그랑지 승수에 대한 미분값 = 0 or 라그랑지 승수 = 0 (05.02 9page)

<=> 람다 = 0 or 람다로 미분한 값 = 0

3) 라그랑지 승수 >= 0

<=> 람다 >= 0



Con)

If) 람다 = 0, 제한조건 없는 상태(제한조건 선과 관련 없이)에서 최적해 찾을 것이고,

elif) 람다 != 0 이라면, 등식 제한조건(제한조건 선 위에서)에서 최적해를 찾을 것

05.03
선형계획법 문제와
이차계획법 문제

05.03 선형계획법 문제와 이차계획법 문제

선형방정식의 최적해 찾기 — 이차형식+선형방정식 최적해 찾기

- 05.01 : 최초 목적함수만 주어졌을 때의 최적화
- 05.02 : 제한조건이 추가되었을 때의 최적화 (등식, 부등식)
- 05.03 : 목적함수, 제한조건이 특정한 경우의 최적화

(선형계획법 문제, 2차계획법 문제)

1) 선형계획법 문제(Linear programming 문제, LP문제)

- 목적함수가 선형함수 인 경우 = 목적함수가 w.Tx 로 계수와 변수를 선형결합 시켜 스칼라 값을 갖는 경우 (5.3.1)

1) 표준형

- 제한조건 2개

Ax = b *A = 등식 제한조건 계수행렬, b = 등식 제한조건 상수 벡터

x >= 0

2) 정규형

- 제한조건 2개

Ax <= b

x >= 0

방정식이나 부등식 제한 조건을 가지는 선형 모형(linear model)의 값을 최소화하는 문제를 **선형계획법(Linear Programming)** 문제라고 한다. LP 문제라고도 한다.

선형계획법 문제의 목적함수는

$$\arg\min_x c^T x \tag{5.3.1}$$

이고 선형 연립방정식으로 된 등식 제한조건

$$Ax = b \tag{5.3.2}$$

과 변수값이 모두 음수가 아니어야하는 부등식 제한조건

$$x \geq 0 \tag{5.3.3}$$

를 동시에 가진다.

선형계획법 문제는 여러가지 형태가 존재하는데 위와 같은 형태를 선형계획법 문제의 기본형(standard form)이라고 한다. 마지막 부등식 제한 조건은 벡터 x의 모든 원소가 양수거나 0이 되어야 한다는 것을 의미한다. 표준형을 확장한 정규형(canonical form) 선형계획법 문제는 부등식 조건을 허용한다.

$$\arg\min_x c^T x \qquad (5.3.1)$$
$$Ax \leq b \qquad (5.3.5)$$
$$x \geq 0 \qquad (5.3.3)$$

- 최적 생산량 예제 (최적화 중 LP문제)

- 해결 방법 : scipy 활용 코드 / CVXPY 활용 코드(직관적이나 속도 느림)

2) 2차계획법 문제(Quadratic programming 문제, QD문제)

- LP 문제 목적함수에 이차형식만 추가됨 (5.3.13)

- 해결 방법 : CVXOPT

방정식이나 부등식 제한 조건을 가지는 일반화된 이차형식(quadratic form)의 값을 최소화하는 문제를 **이차계획법(Quadratic Programming)** 문제라고 한다. QP 문제라고도 한다.

이차계획법 문제의 목적함수는

$$\frac{1}{2}x^T Q x + c^T x \tag{5.3.13}$$

이고 등식 제한조건과 부호 제한조건은 선형계획법 문제와 같다.

$$Ax = b \qquad (5.3.14)$$
$$x \geq 0 \qquad (5.3.15)$$

잔차 제곱합을 최소화하는 예측 모형에 추가적인 제한조건이 있으면 이차계획법 문제가 된다.

06.01 집합	06.01 집합

06.01

집합

06.01 집합

1) 집합과 원소

set이 mutable 자료형

- mutable : 객체 메모리 내용 변경 가능. set~~이 품음~~. (list, dict(key)나 set 자료형의 원소로 사용 불가. (06.01 1page)

frozenset이 immutable 자료형

- immutable : 내용 변경 불가. frozen~~set이 품음~~. (int, float, str, tuple)

· dict의 key : mutable 불가

{ 'A' : 10 o

2 : 10 o

(1,2) : 10 o

[1, 2] : 10 X

사실은 x=1 하면, x → $\boxed{1}$

*파이썬의 모든 변수는 메모리상의 객체의 내용을 가리킴. 즉, 1이라는 숫자가 들어있는 객체의 주소정보를 x가 담고있는 것.

· mutable이라면, x=2 하면 $\boxed{1}$ 이 $\boxed{2}$로 바뀔것

· immutable 이라면, x=2 하면 x → $\boxed{1}$ → $\boxed{2}$

*기존 공간 $\boxed{1}$은 그대로 두고, 주소정보만 바꾸는 것.

2) 집합의 크기 : Cardinality

- Cardinality : 원소의 갯수 (len(A) 로 알 수 있음)

3) 합집합과 교집합

- 합집합 : A.union(A2), A|A2

- 교집합 : A.intersection(A2), A&A2

4) 부분집합

- A.issubset(A2), A<=A2 : A가 A2의 부분집합인지 확인

- A < A2 : A가 진부분집합인지 확인 (진부분집합 : A<B 이지만, A!=B 일 때)

5) 차집합, 여집합

- A - A2, A.difference(A2)

6) 공집합

- A = set([])

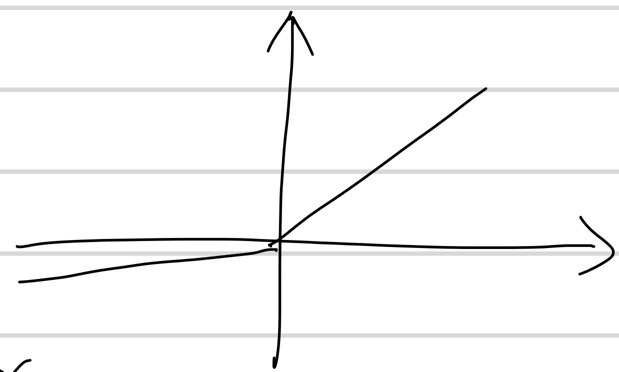
7) 부분집합의 수

- 2**n 개

8) set은 set의 원소가 될 수 없음. frozenset만 set의 원소가 될 수 있음

0425 김도형 강사님 특강

- leaky ReLU



$x \geq 0 \; : \; f(x) = x$

$x < 0 \; : \; f(x) = 0.\square x$

- $\underset{x}{\arg\min} \underbrace{\prod_{i=1}^{4} \exp(x-i)^2} = \underset{x}{\arg\min} \log \prod_{i=1}^{4} e^{(x-i)^2}$

$h(x) = e^{(x-1)^2} \times e^{(x-2)^2} \times e^{(x-3)^2} \times e^{(x-4)^2}$

$= e^{\sum_{i=1}^{4} (x-i)^2}$

$\frac{\partial h}{\partial x} = 0, \quad \underline{\frac{\partial^2 h}{\partial x^2} > 0}.$ $\qquad x = \frac{20}{8} = \frac{5}{2}$

$$\frac{\partial h}{\partial x} = \sum_{i=1}^{4} (x-i)^2 = \begin{array}{l} x^2 - 2x + 1 \\ x^2 - 4x + 4 \\ x^2 - 6x + 9 \\ x^2 - 8x + 16 \end{array} = \frac{4x^2 - 20x}{dx} = \underline{8x - 20}$$