기계학습 4주차 A4 1 Page 정리. 2016104109 컴퓨터공학과 김성수

* Calculus & Optimization

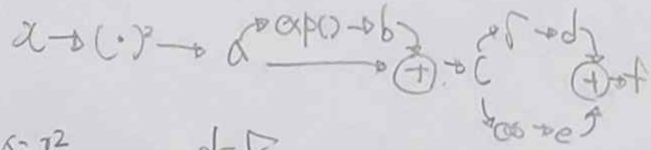
Automatic Differentiation (Back Propagation)

→ Reverse Mode → Foward Mode

$\frac{dy}{dx} = \left(\frac{dy}{db}\cdot\frac{db}{da}\right)\cdot\frac{da}{dx}$ ③ ② ①

$\frac{dy}{dx} = \frac{dy}{db}\cdot\left(\frac{db}{da}\cdot\frac{da}{dx}\right)$ ③ ② ①

$f(x) = \sqrt{x^2+\exp(x^2)} + \cos(x^2+\exp(x^2))$



$a = x^2$ $\quad d = \sqrt{c}$

$b = \exp(a)$ $\quad e = \cos(c)$

$c = a+b$ $\quad f = d+e$

① $\frac{da}{dx} = 2x$ ② $\frac{db}{da} = \exp(a)$ ③ $\frac{\partial d}{\partial c} = \frac{1}{2\sqrt{c}}$

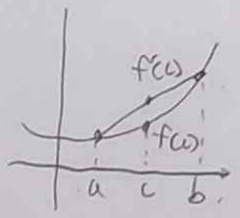
④ $\frac{de}{dc} = -\sin(c)$ ⑤ $\frac{df}{dc} = \underset{③}{\frac{dd}{dc}} + \underset{④}{\frac{de}{dc}}$

$\therefore \frac{df}{da} = \frac{df}{dc}\cdot\frac{dc}{da} = \frac{df}{dc}\cdot\left(\frac{db}{da}+1\right)$

$\frac{df}{dx} = \frac{df}{da}\cdot\frac{da}{dx}$

* Convex Function

$f(\theta x + (1-\theta)y) \le \theta f(x) + (1-\theta) f(y)$



함수값의 대응집 $a \le c \le b$의 all c

내분점의 함수값 $f(c)$.

$\le$

$f(c) \le f'(c)$의 경우임.

+ 직선 방정식에서 법선벡터 $ax+bx_2+c=0$, $n = \alpha\begin{bmatrix}a\\b\end{bmatrix}$

+ 평면 〃 〃 $x_1+bx_2+cx_3+d=0$, $n = \alpha\begin{bmatrix}a\\b\\c\end{bmatrix}$

* Equality Constraints.

$\nabla g(x)$, $\nabla f(x)$ ⇒ Parallel at x^*.

* Lagrangian Function

⇒ $L(x,\lambda) = f(x) + \lambda g(x)$.

① $\nabla_x L(x,\lambda) = 0 \Leftrightarrow \nabla_x f(x) + \lambda \nabla_x g(x) = 0$.

② $\nabla_\lambda L(x,\lambda) = 0 \Leftrightarrow g(x) = 0$.

* Inequality Constraints.

$\nabla_x L(x,\lambda) = 0$.

$g(x^*) \le 0$.

$\lambda \ge 0$

$\lambda g(x^*) = 0$

⇒ KKT condition,

* p.14 Example 참고

* Multiple Constraints

$$L(x,\lambda,v) = f_0(x) + \sum_{i=1}^{m}\lambda_i f_i(x) + \sum_{i=1}^{k} v_i h_i(x).$$

$$= f_0(x) + \lambda^T f(x) + v^T h(x).$$

cf) $\min f(x_1,x_2)$ s.t $g(x_1,x_2) = 0$