행렬곱의 Backpropagation

조희철

2020년 12월 10일

$m \times n$ 행렬 X, $n \times k$ 행렬 W와 함수 $f : \mathbb{R}^{mk} \to \mathbb{R}$에 대하여, X, W의 각 원소에 대한 미분을 구해보자.

$$\begin{matrix} \boxed{X} & & & & \\ & \longmapsto & Y = XW & \overset{f}{\longmapsto} & f(Y) = f(XW) \in \mathbb{R} \\ \boxed{W} & & & & \end{matrix}$$

행렬 Y는 함수 f를 거쳐 최종적으로 loss(cost) 값인 scalar가 된다.

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{pmatrix}, \; W = \begin{pmatrix} w_{11} & \cdots & w_{1k} \\ \vdots & \ddots & \vdots \\ w_{n1} & \cdots & w_{nk} \end{pmatrix}, \; Y = XW = \begin{pmatrix} y_{11} & \cdots & y_{1k} \\ \vdots & \ddots & \vdots \\ y_{m1} & \cdots & y_{mk} \end{pmatrix}.$$

$\frac{\partial f}{\partial Y}$(Jacobian)를 다음과 같이 정의하면(또는 주어졌다고 하면),

$$\frac{\partial f}{\partial Y} := \begin{pmatrix} \frac{\partial f}{\partial y_{11}} & \cdots & \frac{\partial f}{\partial y_{1k}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f}{\partial y_{m1}} & \cdots & \frac{\partial f}{\partial y_{mk}} \end{pmatrix} \quad \leftarrow m \times k$$

X, W의 각 원소별 미분(Jacobian)은 다음과 같다.

$$\begin{pmatrix} \frac{\partial f}{\partial x_{11}} & \cdots & \frac{\partial f}{\partial x_{1n}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f}{\partial x_{m1}} & \cdots & \frac{\partial f}{\partial x_{mn}} \end{pmatrix} = \underbrace{\frac{\partial f}{\partial Y} W^T}_{(m\times k)(k\times n)} \quad \leftarrow m \times n \tag{1}$$

$$\begin{pmatrix} \frac{\partial f}{\partial w_{11}} & \cdots & \frac{\partial f}{\partial w_{1k}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f}{\partial w_{n1}} & \cdots & \frac{\partial f}{\partial w_{nk}} \end{pmatrix} = \underbrace{X^T \frac{\partial f}{\partial Y}}_{(n\times m)(m\times k)} \quad \leftarrow n \times k \tag{2}$$

증명을 하기에 앞서 문제를 다시 한번 정리해 보자. X, W가 주어져 있고, $Y = XW, f(Y) \in \mathbb{R}$로 정의된다고 하자. 이 때 f를 Y의 각 원소별로 미분한 Jacobian $\frac{\partial f}{\partial Y}$가 주어진다고 하면, f를 X, W의 각 원소로 미분한 Jacobian은 각각 식(1), (2)가 된다는 것이다.

식(1), (2) 중에 하나만 증명하면, 나머지 하나는 대칭성에 의해 증명된다. 그래서 우리는 식(1)만 증명하고자 한다. 식(1)을 3단계로 나누어서 증명해보자.

- X, W가 각각 $1 \times n$, $n \times 1$인 경우:

$$XW = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = x_1 w_1 + x_2 w_2 + \cdots + x_n w_n = y = Y$$

$$
\begin{aligned}
\frac{\partial f}{\partial x_i} &= \frac{\partial f}{\partial y}\frac{\partial y}{\partial x_i} = \frac{\partial f}{\partial y} w_i \\
\frac{\partial f}{\partial X} &= \begin{pmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \cdots & \frac{\partial f}{\partial x_n} \end{pmatrix} \\
&= \begin{pmatrix} \frac{\partial f}{\partial y} w_1 & \frac{\partial f}{\partial y} w_2 & \cdots & \frac{\partial f}{\partial y} w_n \end{pmatrix} \\
&= \frac{\partial f}{\partial y} \begin{pmatrix} w_1 & w_2 & \cdots & w_n \end{pmatrix} \\
&= \frac{\partial f}{\partial y} \begin{pmatrix} w_1 & w_2 & \cdots w_n \end{pmatrix} \\
&= \underbrace{\frac{\partial f}{\partial Y} W^T}_{(1\times 1)(1\times n)} \quad \leftarrow 1 \times n
\end{aligned} \tag{3}
$$

- X, W 가 각각 $1 \times n, n \times k$ 인 경우:

$$
\begin{aligned}
XW &= \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix} \begin{pmatrix} w_{11} & \cdots & w_{1k} \\ \vdots & \ddots & \vdots \\ w_{n1} & \cdots & w_{nk} \end{pmatrix} \\
&= \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix} \begin{pmatrix} \boxed{W_1} & \boxed{W_2} & \cdots & \boxed{W_k} \end{pmatrix} \quad \leftarrow W_i \text{는 } W \text{의 열벡터} \\
&= \begin{pmatrix} y_1 & y_2 & \cdots & y_k \end{pmatrix} = Y
\end{aligned}
$$

먼저, chain rule에 의해,

$$
\frac{\partial f}{\partial x_i} = \frac{\partial f}{\partial y_1}\frac{\partial y_1}{\partial x_i} + \frac{\partial f}{\partial y_2}\frac{\partial y_2}{\partial x_i} + \cdots + \frac{\partial f}{\partial y_k}\frac{\partial y_k}{\partial x_i} \quad \text{by chain rule}
$$

가 되고, $\frac{\partial f}{\partial X} = \begin{pmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \cdots & \frac{\partial f}{\partial x_n} \end{pmatrix}$, $\frac{\partial y_i}{\partial X} = \begin{pmatrix} \frac{\partial y_i}{\partial x_1} & \frac{\partial y_i}{\partial x_2} & \cdots & \frac{\partial y_i}{\partial x_n} \end{pmatrix}$ 이므로,

$$
\begin{aligned}
\frac{\partial f}{\partial X} &= \frac{\partial f}{\partial y_1}\frac{\partial y_1}{\partial X} + \frac{\partial f}{\partial y_2}\frac{\partial y_2}{\partial X} + \cdots + \frac{\partial f}{\partial y_k}\frac{\partial y_k}{\partial X} \quad \text{by chain rule} \\
&= \frac{\partial f}{\partial y_1} W_1^T + \frac{\partial f}{\partial y_2} W_2^T + \cdots + \frac{\partial f}{\partial y_k} W_k^T \quad \text{by (3)} \\
&= \begin{pmatrix} \frac{\partial f}{\partial y_1} & \frac{\partial f}{\partial y_2} & \cdots & \frac{\partial f}{\partial y_k} \end{pmatrix} \begin{pmatrix} \boxed{W_1^T} \\ \boxed{W_2^T} \\ \vdots \\ \boxed{W_k^T} \end{pmatrix} \quad \leftarrow \text{(행벡터의 일차결합)} \\
&= \begin{pmatrix} \frac{\partial f}{\partial y_1} & \frac{\partial f}{\partial y_2} & \cdots & \frac{\partial f}{\partial y_k} \end{pmatrix} W^T \\
&= \underbrace{\frac{\partial f}{\partial Y} W^T}_{(1\times k)(k\times n)} \quad \leftarrow 1 \times n
\end{aligned}
$$

- X, W 가 각각 $m \times n, n \times k$ 인 경우: X 의 각 행을 서로 독립적이다. 그렇기 때문에, X 의 각 행은 Y 의 각 행에 대응된다. 따라서 행으로의 확장은 X, $\frac{\partial f}{\partial Y}$ 를 행으로 확장하면 된다.

$$
\frac{\partial f}{\partial X} = \underbrace{\frac{\partial f}{\partial Y} W^T}_{(m\times k)(k\times n)} \quad \leftarrow m \times n
$$