MŰVELETEK VEKTOROKKAL ÉS MÁTRIXOKKAL

1. Bevezetés

A vektorok és a mátrixok a lineáris algebrából jól ismert egy-, illetve kétdimenziós elemek. Itt megvizsgáljuk néhány alkalmazásukat.

A vektorokat a szokásos módon jelöljük, sokszor helytakarékosság miatt a transzponáltjukat is felhasználva:

$$a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = [a_1, a_2, \ldots, a_n]^T$$

A mátrixokat nagybetűvel jelöljük, és felbonthatjuk sorok és oszlopokra:

$$B = B_{m \times n} = \begin{bmatrix} b^1 \\ b^2 \\ \vdots \\ b^m \end{bmatrix} = \begin{bmatrix} b_1 & b_2 & \ldots & b_n \end{bmatrix}$$

Jól látszik, hogy a sorvektorokat felső, az oszlopvektorokat alsó indexszel szokás jelölni. Sorvektorra szokás még a transzponált jel alkalmazása, de ez csak akkor alkalmazható, ha az oszlopvektorokat nem használjuk, hiszen egyébként a két jelölés összekeveredik:

$$C = C_{m \times n} = \begin{bmatrix} c_1^T \\ c_2^T \\ \vdots \\ c_m^T \end{bmatrix}$$

2. Deriválás

Tegyük fel, hogy n darab változónk van, melyeket $x_1, x_2, \ldots, x_n$-nel jelölünk, és beleraktuk az x vektorba: $x = [x_1, x_2, \ldots, x_n]^T$

Első feladatként nézzük meg, hogy hogyan lehet $x^T x$ deriváltját képezni:

$$\frac{d(x^T x)}{dx} = 2x$$

Bizonyítás: Először is tisztázzuk, mit jelent vektor szerinti deriválás:

$$\frac{dJ}{dx} \triangleq \begin{bmatrix} \frac{\partial J}{x_1} \\ \frac{\partial J}{x_2} \\ \vdots \\ \frac{\partial J}{x_n} \end{bmatrix}$$

Tudjuk, hogy $x^T x = x_1^2 + x_2^2 + \cdots + x_n^2$. $\frac{\partial x^T x}{\partial x_i} = 2x_i$.
Ezért

$$\frac{dx^T x}{x} = \begin{bmatrix} 2x_1 \\ 2x_2 \\ \vdots \\ 2x_n \end{bmatrix} = 2x$$

Ezzel az állítást bebizonyítottuk.
Vegyünk egy új tétel, amelyik
Tétel:

$$\frac{dz^T x}{dx} = \frac{dx^T z}{dx} = z$$

Bizonyítás: Fejtsük ki
Tétel:

$$\frac{d(x^T Ax)}{dx} = A^T x + Ax$$

Bizonyítás: Írjuk fel az A mátroxot sorvektorok segítségével, és szorozzuk meg jobbról x vektorral (az eredmény egy n elemű vektor lesz:

$$A_{n\times n}x = \begin{bmatrix} a^1 \\ a^2 \\ \vdots \\ a^n \end{bmatrix} x = \begin{bmatrix} a^1 x \\ a^2 x \\ \vdots \\ a^n x \end{bmatrix}$$

Most pedig az eredményt balról szorozzuk x^T-tal:

$$x^T Ax = x_1 a^1 x + x_2 a^2 x + \cdots + x_n a^n x$$

Ezt most deriváljuk le x_1 szerint:

$$\frac{\partial x^T Ax}{\partial x_1} = 2a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n +$$

$$+a_{21}x_2 + \cdots + a_{n1}x_n$$

Mivel az első tagot ketté tudjuk bontani ($2a_{11}x_1 = a_{11}x_1 + a_{11}x_1$), az összeget fel tudjuk írni:

$$\frac{\partial x^T Ax}{\partial x_1} = a^1 x + x^T a_1 = a^1 x + a_1^T x$$

(Figyelem! Itt a_1^Taz első oszlopvektor transzponáltját jelenti!)

Az i-edik parciális derivált ehhez teljesen hasonlóan kapható:

$$\frac{\partial x^T A x}{\partial x_i} = a^i x + a_i^T x$$

Az összes parciális deriváltat pedig összerakhatjuk:

$$\frac{\partial x^T A x}{\partial x} = Ax + A^T x$$

Fontos megjegyzés: A tétel egyik speciális esete, hogy egy szimmetrikus B mátrixra igaz, hogy

$$\frac{d(x^T B x)}{dx} = 2Bx$$

hiszen $B^T = B$.