Differentiate Matrix Formula 線形代数式を微分する

ryamada

2016年12月27日

1 最小二乗法

$$y \sim X\mathbf{a}$$

なる関係式があり、Xの行数が列数より多いとき、

$$\sum_{i=1}^{n}(y_i - \hat{y_i})^2 = (y - X\mathbf{a})^T \cdot (y - X\mathbf{a})$$

を最小にするような$\hat{\mathbf{a}}$が、推定値であった。

そして

$$\hat{\mathbf{a}} = (X^T X)^{-1} X^T y$$

で求まるのであった。

2 偏微分方程式

$$f(\mathbf{a}) = f((a_1, \ldots, a_m)) = \sum_{i=1}^{n}(y_i - \hat{y_i})^2 = (y - X\mathbf{a})^T \cdot (y - X\mathbf{a})$$

は、y,Xが与えられているとき、m個の変数$(a_1, \ldots, a_m)$の二次式となっているスカラー関数である。

今、この$f(\mathbf{a})$を最小にする$\mathbf{a} = (a_1, \ldots, a_m)$とは、

$$\frac{\partial f(\mathbf{a})}{\partial a_i} = 0$$

がすべてのa_iについて成り立つ場合である。

スカラーを返す$z^T w$は$z^T w = w z^T$であることを使うと

$$f(\mathbf{a}) = (y - X\mathbf{a})^T \cdot (y - X\mathbf{a}) = y^T y - 2(X\mathbf{a})^T y + (X\mathbf{a})^T \cdot (X\mathbf{a})$$

さらに変形して

$$f(\mathbf{a}) = y^T y - 2(X^T y) \cdot \mathbf{a} + \mathbf{a}^T X^T X\mathbf{a}$$

これを$\mathbf{a}$で偏微分する。

第1項は0

第2項は、各成分に$-2y^T X$の各成分が残る。

第３項は、$X^T X = Z$と置くと

$$\mathbf{a}^T Z\mathbf{a} = \sum_{i=1}^{m}\sum_{j=1}^{m} a_i a_j Z_{ij}$$

であり、そのa_kによる偏微分は

$$\frac{\partial}{\partial a_k}(\sum_{i=1}^{m}\sum_{j=1}^{m} a_i a_j Z_{ij}) = 2\sum_{i=1}^{m} Z_{ik} a_i$$

となり、$k = 1, \ldots, m$について合わせると

$$2Z\mathbf{a}$$

となるから、 結局、

$$0 - 2X^T y + 2Z\mathbf{a} = 2(X^T X)\mathbf{a} - 2X^T y = \mathbf{0}$$

となり、

$$(X^T X)\mathbf{a} = X^T y$$

が得られる。

ここから、$X^T X$に逆行列があるときは

$$\mathbf{a} = (X^T X)^{-1} X^T y$$

が得られる。