5.27

5.5.5節の ξ はスカラーであるが、
この問では ξ はベクトルである。D次元とする。

多変数のテイラー展開の式。x をD次元として

$$f(x) = f(a) + \sum_{i=0}^{D} \frac{\partial f(a)}{\partial x_i}(x_i - a_i) + \frac{1}{2!}\sum_i \sum_j \frac{\partial^2 f(a)}{\partial x_i \partial x_j}(x_i - a_i)(x_j - a_j) + \frac{1}{3!}\sum_i \sum_j \sum_k \frac{\partial^3 f(a)}{\partial x_i \partial x_j \partial x_k}(x_i - a_i)(x_j - a_j)(x_k - a_k) \cdots$$

ここで

$$\sum_i \frac{\partial f(a)}{\partial x_i}(x_i - a_i) = \nabla f(a)^T (x - a)$$

$$\sum_i \sum_j \frac{\partial^2 f(a)}{\partial x_i \partial x_j}(x_i - a_i)(x_j - a_j) = (x-a)^T \nabla\nabla f(a) (x-a)$$

なので

$$f(x) = f(a) + \nabla f(a)^T (x-a) + \frac{1}{2!}(x-a)^T \nabla\nabla f(a)(x-a) + O(|x-a|^3)$$

$y(s(x,\xi))$ を $\xi = 0$ のまわりでテイラー展開する

$$y(s(x,\xi)) = y(s(x,\xi))\Big|_{\xi=0} + \nabla_\xi y(s(x,\xi))\Big|_{\xi=0} + \frac{1}{2!}\xi^T \nabla_\xi \nabla_\xi y(s(x,\xi))\Big|_{\xi=0} \xi + O(|\xi|^3)$$

ここで $s(x,\xi) = x + \xi$ とすると

$$y(s(x,\xi))\Big|_{\xi=0} = y(x) \quad \cdots ①$$

次に $\nabla_\xi y(s(x,\xi))\big|_{\xi=0}$ について　　$s=x+\xi$ なので $\frac{\partial s_j}{\partial \xi_i}=\delta_{ji}$

成分を考えると

$$\frac{\partial}{\partial \xi_i} y(s(x,\xi))\Big|_{\xi=0} = \sum_{j=0}^{D} \frac{\partial y}{\partial s_j}\frac{\partial s_j}{\partial \xi_i}\Big|_{\xi=0} = \frac{\partial y}{\partial s_i}\Big|_{\xi=0} \cdots ②$$

(例) $y=s^2=s_1^2+s_2^2$ のとき

$\frac{\partial y}{\partial s_1}\Big|_{\xi=0}=2s_1\Big|_{\xi=0}=2(x_1+\xi_1)_{\xi=0}=2x_1=\frac{\partial y(x)}{\partial x_1}$

よって $\nabla_\xi y(s(x,\xi))\big|_{\xi=0} = \nabla_s y(s)\big|_{\xi=0}$

又 $s=x+\xi$ なので $\nabla_s y(s)\big|_{\xi=0} = \nabla_x y(x)$

よって

$$\nabla_\xi y(s(x,\xi))\big|_{\xi=0} = \nabla_x y(x) \cdots ③$$

次に $\nabla_\xi \nabla_\xi y(s(x,\xi))\big|_{\xi=0}$ を考える

$$ij\text{成分} = \frac{\partial^2 y}{\partial \xi_i \partial \xi_j}\Big|_{\xi=0} = \frac{\partial}{\partial \xi_i}\frac{\partial y}{\partial \xi_j}\Big|_{\xi=0} \overset{②より}{=} \frac{\partial}{\partial \xi_i}\frac{\partial y}{\partial s_j}\Big|_{\xi=0} = \frac{\partial\partial y}{\partial s_j \partial \xi_i}\Big|_{\xi=0} \overset{②より}{=} \frac{\partial^2 y}{\partial s_j \partial s_i}\Big|_{\xi=0}$$

なので $\nabla_\xi \nabla_\xi y(s(x,\xi))\big|_{\xi=0} = \nabla_s \nabla_s y(s(x,\xi))\big|_{\xi=0}$ である。

ここで $\nabla_s \nabla_s y(s)\big|_{\xi=0} = \nabla_x \nabla_x y(x)$ なので　　(例) $y=s_1^3+s_2^3$ のとき $\frac{\partial^2 y}{\partial s_1^2}\Big|_{\xi=0}=6s_1\Big|_{\xi=0}=6x=\frac{\partial^2 y(x)}{\partial x^2}$

$$\nabla_\xi \nabla_\xi y(s(x,\xi))\big|_{\xi=0} = \nabla_x \nabla_x y(x) \cdots ④$$

①, ③, ④ を使うと y のテーラー展開はこのようになる。

$$y(s(x,\xi)) = y(x) + \nabla y(x) + \frac{1}{2!}\xi^T \nabla\nabla y(x)\xi + O(|\xi|^3)$$

yのテーラ展開を $\tilde{E}$ の式に入れる

$$\tilde{E} = \frac{1}{2}\iiint \{y(s(x,\xi)) - t\}^2 p(t|x)\,p(x)\,p(\xi)\,dx\,dt\,d\xi$$

$$= \frac{1}{2}\iiint \{\underbrace{y(x) + \nabla y\,\xi + \frac{1}{2!}\xi^T \nabla\nabla y\,\xi + O(|\xi|^3) - t}_{⑤}\}^2 p(t|x)\,p(x)\,p(\xi)\,dx\,dt\,d\xi$$

$$⑤ = \{y(x)-t\}^2 + \underbrace{(\nabla y\,\xi)^2}_{= \xi^T \nabla y \nabla y^T \xi} + 2\{y(x)-t\}\nabla y\,\xi + \{y(x)-t\}\xi^T \nabla\nabla y\,\xi + O(|\xi|^3)$$

よって

$$\tilde{E} = \frac{1}{2}\iiint \{y(x)-t\}^2 p(t|x)\,p(x)\,p(\xi)\,dx\,dt\,d\xi \quad \text{---}⑥$$

$$+ \iiint \{y(t)-t\}\nabla y\,\xi\, p(t|x)\,p(x)\,p(\xi)\,dx\,dt\,d\xi \quad \text{---}⑦$$

$$+ \frac{1}{2}\iiint \xi^T[\nabla y \nabla y^T + \{y(x)-t\}\nabla\nabla y]\xi\, p(t|x)\,p(x)\,p(\xi)\,dx\,dt\,d\xi \quad \text{---}⑧$$

$$+ \frac{1}{2}\iiint O(|\xi|^3)\, p(t|x)\,p(x)\,p(\xi)\,dx\,dt\,d\xi \quad \text{---}⑨$$

$\int p(\xi)\,d\xi = 1$ なので

$$⑥ = \frac{1}{2}\iint \{y(x)-t\}^2 p(t|x)\,p(x)\,dx\,dt = E(x)$$

$\int \xi\, p(\xi)\,d\xi = 0$ なので ⑦ $= 0$

p268で ξ は平均0で十分小さい分散と仮定してるので $|\xi|^3$ の期待値が $O(|\xi|^3)$ 程度と思って良い

$\int O(|\xi|^3)\,p(\xi)\,d\xi = \mathbb{E}(O(|\xi|^3)) = O(|\xi|^3)$ なので ⑨ $= O(|\xi|^3)$

よって $O(|\xi|^3)$ を無視して

$$\tilde{E}(x) = E(x) + \frac{1}{2}\iiint \xi^T[\nabla y \nabla y^T + \{y(x)-t\}\nabla\nabla y]\xi\, p(t|x)\,p(x)\,p(\xi)\,dx\,dt\,d\xi$$

← tについて積分した

$$= E(x) + \frac{1}{2}\iint \xi^T[\nabla y \nabla y^T + \{y(x)-\mathbb{E}(t|x)\}\nabla\nabla y]\xi\, p(x)\,p(\xi)\,dx\,d\xi$$

を得る。

ここで $y = \mathbb{E}(t|x) + O(\xi^2)$ … (5.133) なので

$$\tilde{E} = E(x) + \frac{1}{2}\iint \xi^T [\nabla y \nabla y^T + O(\xi^2)]\xi p(x)p(\xi)dxd\xi$$

ここで $\int \xi^T O(\xi^2)\xi p(\xi)d\xi = \mathbb{E}(O(|\xi|^4)) = O(|\xi|^4)$ は無視できる

2次元の場合

$$\int (\xi_1\ \xi_2)\begin{pmatrix} O(\xi^2)_{11} & O(\xi^2)_{12} \\ O(\xi^2)_{21} & O(\xi^2)_{22}\end{pmatrix}\begin{pmatrix}\xi_1 \\ \xi_2\end{pmatrix} p(\xi)d\xi_1 d\xi_2 = \int O(|\xi|^4)p(\xi)d\xi$$

なので

$$\tilde{E} = E(x) + \frac{1}{2}\iint \xi^T \nabla y \nabla y^T \xi p(x)p(\xi)dxd\xi$$

を得る

ここで $\xi^T\nabla y \nabla y^T \xi \overset{\text{スカラーなので}}{=} \nabla y^T \xi\xi^T \nabla y$ なので

$$\tilde{E} = E(x) + \frac{1}{2}\int \nabla y^T \int \xi\xi^T p(\xi)d\xi \nabla y\, p(x)dx$$

ここで問題文より $p(\xi)$ は平均0で十分小さい等方的な分散を持つので

$$p(\xi) = N(\xi|0, \alpha I) = \frac{1}{(2\pi)^{D/2}}\exp\left(-\frac{1}{2}\xi^T\alpha I\xi\right) = \frac{1}{(2\pi)^{D/2}}\exp\left(-\frac{\alpha}{2}|\xi|^2\right)$$

$$= \underbrace{\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{\alpha}{2}\xi_1^2\right)}_{=N(\xi_1|0,\alpha)}\underbrace{\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{\alpha}{2}\xi_2^2\right)}_{=N(\xi_2|0,\alpha)}\cdots = \prod_i^D N(\xi_i|0,\alpha)$$

なので

$$\int \xi\xi^T p(\xi)d\xi = \iint \begin{pmatrix}\xi_1^2 & \xi_1\xi_2 \\ \xi_1\xi_2 & \xi_2^2\end{pmatrix} N(\xi_1|0,\alpha)N(\xi_2|0,\alpha)d\xi_1 d\xi_2$$

(1,1) 成分 $= \mathbb{E}(\xi_1^2) = \alpha \leftarrow \mathbb{E}(\xi_i^2) = N(\xi_i|0,\alpha)$ の分散 $= \alpha$

(1,2) 〃 $= \mathbb{E}(\xi_1)\mathbb{E}(\xi_2) = 0 \leftarrow \mathbb{E}(\xi_i) = N(\xi_i|0,\alpha)$ の平均 $= 0$

(2,1) 〃 〃 $= 0$

(2,2) 〃 $= \mathbb{E}(\xi_2^2) = \alpha$

よって

$\int \xi\xi^T p(\xi)d\xi = \alpha I$ である。

よって

$$\tilde{E} = E(x) + \frac{1}{2}\int \nabla y^T \alpha I \nabla y \, p(x)\, dx$$

$$= E(x) + \frac{\alpha}{2}\int \nabla y^T \nabla y \, p(x)\, dx$$

$$= E(x) + \alpha \frac{1}{2}\int |\nabla y|^2 p(x)\, dx$$

$\lambda = \alpha$, $\Omega = \frac{1}{2}\int |\nabla y|^2 p(x)\, dx$ …(5.137) を得る。