Machine Learning

Assignment 03: Probability Review

Übung 01

Beweis: $Var(X) = E(X^2) - (E(X)^2)$
Vorgabe: $\mu = E(X)$

$$Var(X) = E((X-\mu)^2) = E\left(\left(X - E(X)\right)^2\right) = E\left(X^2 - 2 * X * E(X) + \left(E(X)\right)^2\right)$$
$$Var(X) = E(X^2) - 2E\left(X * E(X)\right) + E\left(\left(E(X)\right)^2\right)$$
$$Var(X) = E(X^2) - 2 * E(X) * E(X) + E\left(\left(E(X)\right)^2\right)$$
$$Var(X) = E(X^2) - 2 * \left(E(X)\right)^2 + \left(E(X)\right)^2$$
$$Var(X) = E(X^2) - \left(E(X)\right)^2$$

Übung 02

Beweis: $Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$
Vorgabe: $\mu_X = E(X)$, $\mu_Y = E(Y)$, $Var(X) = E((X-\mu_X)^2)$

$$Var(X+Y) = E((X+Y-\mu_X-\mu_Y)^2) = E\left(\left((X-\mu_X)(Y-\mu_Y)\right)^2\right)$$
$$Var(X+Y) = E((X-\mu_X)^2 + 2(X-\mu_X)(Y-\mu_Y) + (Y-\mu_Y)^2)$$
$$Var(X+Y) = E((X-\mu_X)^2) + E((Y-\mu_Y)^2) + E\left(2(X-\mu_X)(Y-\mu_Y)\right)$$
$$Var(X+Y) = Var(X) + Var(Y) + 2E\left((X-\mu_X)(Y-\mu_Y)\right)$$
$$Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$$

$$E\left((X-\mu_X)(Y-\mu_Y)\right) = E(XY - X\mu_Y - \mu_X Y + \mu_X\mu_Y) = E(XY) - E(X\mu_Y) - E(\mu_X Y) + E(\mu_X\mu_Y)$$
$$E\left((X-\mu_X)(Y-\mu_Y)\right) = E(XY) - \mu_Y E(X) - \mu_X E(Y) + \mu_X\mu_Y E(1)$$
$$E\left((X-\mu_X)(Y-\mu_Y)\right) = E(XY) - E(Y) * E(X) - E(X) * E(Y) + E(X) * E(Y)$$
$$E\left((X-\mu_X)(Y-\mu_Y)\right) = E(XY) - E(X) * E(Y) = Cov(X,Y)$$

Übung 03

Beweis: Kovarianz Matrix ist immer symmetrisch und positiv semi-definiert.

Vorgabe: $C = E((x-\bar{x})(x-\bar{x})^T)$, beliebiger *Vektor* $v \in R^n$

$$v^T C v = v^T E\left((x-\bar{x})(x-\bar{x})^T\right) v = E(v^T(x-\bar{x})(x-\bar{x})^T v) = E(s^2) = {\sigma_s}^2$$
$$\sigma_s = v^T(x-\bar{x}) = (x-\bar{x})^T v$$
$$=> \; v^T C v = {\sigma_s}^2 \geq 0 \qquad\qquad qed.$$

Aus der Definition der Kovarianz-Matrix: $\sum_{i,j} = E_{x\sim p}\left[\left(x_i - \mu_j\right)\left(x_i - \mu_j\right)\right]$
geht hervor, dass die Reihenfolge keine Rolle spielt, also dass $\sum_{i,j} = \sum_{j,i}$ *ist, und damit ist die Kovarianz-Matrix symmetrisch.*

Marcel Zauder 16-124-856

Machine Learning

Assignment 03: Probability Review

Übung 04

$f(x) = \{ \begin{matrix} {}^{1}/_{b-a} & ,a \leq x \leq b \\ 0 & ,sonst \end{matrix}$

$1 = \int_a^b {}^{1}/_{b-a}\, dx = \left[\frac{x}{b-a}\right]_a^b = \frac{b-a}{b-a} = 1$ *qed.*

Übung 05

$f(x) = \{ \begin{matrix} \lambda e^{-\lambda x} & ,0 \leq x \leq \infty \\ 0 & ,sonst \end{matrix}$

$1 = \lim_{h\to\infty} \int_0^h \lambda e^{-\lambda x}\, dx = \lim_{h\to\infty} \left[-e^{-\lambda x}\right]_0^h = \lim_{h\to\infty} -e^{-\lambda h} - (-e^{-\lambda * 0}) = 0 - (-1) = 1$ *qed.*

Übung 06

Seien $X_1, X_2, \cdots, X_n$ beliebige unabhängige und identische, Poisson-verteilte Zufallsvariablen mit $P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$.

<u>Gesucht:</u> Das λ, das die Wahrscheinlichkeit von $X_1, X_2, \cdots, X_n$ maximiert.

$$L(p|x_1, \cdots, x_n) = \prod_{i=1}^n \frac{\lambda^{x_i}}{x_i!} e^{-\lambda}$$
$$L(p|x_1, \cdots, x_n) = \ln\left(\prod_{i=1}^n \frac{\lambda^{x_i}}{x_i!} e^{-\lambda}\right) = \sum_{i=1}^n \ln\left(\frac{\lambda^{x_i}}{x_i!} e^{-\lambda}\right) = \sum_{i=1}^n \ln(\lambda^{x_i}) - \ln(x_i!) + \ln\left(e^{-\lambda}\right)$$
$$L(p|x_1, \cdots, x_n) = \sum_{i=1}^n x_i * \ln(\lambda) - \ln(x_i!) - \lambda = \ln(y) * \sum_{i=1}^n x_i - \sum_{i=1}^n \ln(x_i!) - \sum_{i=1}^n \lambda$$
$$L(p|x_1, \cdots, x_n) = \ln(\lambda) * n * \bar{x} - \sum_{i=1}^n \ln(x_i!) - n * \lambda \qquad mit\ \bar{x} = Durchschnitt\ aller\ x_i$$

$$\frac{dL}{d\lambda} = {}^{1}/_{\lambda} * n * \bar{x} - n \qquad \frac{dL}{dp} = 0$$

$$\begin{aligned} 0 &= {}^{1}/_{\lambda} * n * \bar{x} - n && |\ /n \quad mit\ n \neq 0\ und\ \lambda > 0 \\ 0 &= {}^{1}/_{\lambda} * \bar{x} - 1 && | + 1 \\ 1 &= {}^{1}/_{\lambda} * \bar{x} && | * \lambda \\ \lambda &= \bar{x} \end{aligned}$$

$$\frac{d^2L}{d\lambda d\lambda} = -\frac{1}{\lambda^2} * n * \bar{x} < 0 \qquad \forall\lambda, da\ n > 0\ und\ x_i \epsilon N_0\ \forall i \in \{1, \cdots, n\}$$

L hat somit ein globales Maximum an der Stelle $\lambda = \bar{x} = mean(x_1, \ldots, x_n)$.

Machine Learning

Assignment 03: Probability Review

Übung 07

Vorgabe: $E(X) = 0\,, Cov(x) = Var(X) = I\,, E(Y) = \mu\,, Cov(Y) = Var(Y) = \sigma I$

Gesucht: $E(Z) und\ Var(Z), wobei\ Z = AX + Y$

$$E(Z) = E(AX + Y) = E(AX) + E(Y) = A * E(X) + E(Y) = A * 0 + \mu = \mu$$

$$Var(Z) = Var(AX + Y) = A^2 * Var(X) + Var(Y) + 2A * Cov(X, Y) = A^2 * I + \sigma I = A^2 + \sigma * I$$

, da X und Y unabhängig voneinander sind ist $Cov(X, Y) = 0$.

Übung 08

$$X, Y \sim U[0,1]\ mit\ E(\min(X, Y)) = \frac{1}{2} * E(X + Y - |X - Y|)$$

$$E(\min(X, Y)) = \frac{1}{2} * (E(X) + E(Y) - E(|X - Y|) = \frac{1}{2} * (\frac{1}{2} + \frac{1}{2} - E(|X - Y|)$$

$$E(\min(X, Y)) = \frac{1}{2} * \left(1 - E(|X - Y|)\right) = \frac{1}{2} - \frac{1}{2} * E(|X - Y|) = \frac{1}{2} - \frac{1}{2} E(Z) \quad mit\ Z = |X - Y|$$

$$g(X, Y) = |X - Y| = \{ \begin{matrix} X - Y, & X \geq Y \\ Y - X, & X < Y \end{matrix}$$

$$f_{x,y}(X, Y) = f_x(X) * f_y(Y) \qquad , da\ X\ und\ Y\ unabhängig\ sind.$$

$$f_x(X) = f_y(Y) = \{ \begin{matrix} 1, & x \in [0,1] \\ 0, & sonst \end{matrix}$$

$$f_{x,y}(X, Y) = 1\ in\ [0,1]x[0,1]$$

$$=> E(Z) = E\big(g(X, Y)\big)$$

$$E(Z) = \int_0^1 \int_0^1 g(X, Y) f_{x,y}(X, Y) dX dY = \int_0^1 \int_0^1 |X - Y| dX dY$$

$$E(Z) = \int_0^1 \int_0^x (X - Y) dY dX + \int_0^1 \int_x^1 (Y - X) dY dX = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

⇨ $\frac{1}{2} - \frac{1}{2} * E(Z) = \frac{1}{2} - \frac{1}{6} = \frac{1}{3}$