Malte Schwarzer (340819)
Nils Barge (330849)
Bianca Lüders (327210)

Mittwoch, 21. Oktober 2015

Exercise Sheet 1 - Math Primer

1.1 Distributions and expected values

a) Die Fläche unter der gegebenen Funktion ist 2:

$$A = \int_0^\pi \sin(x)\,dx = [-\cos(x)]_0^\pi = -\cos(\pi) + \cos(0) = 2$$

Um diese auf 1 zu skalieren (die Fläche unter der Kurve einer Wahrscheinlichkeitsverteilung muss 1 sein), muss $c = \frac{1}{2}$ sein.

Der analytische Nachweis hierfür ist:

$$0 \le x \le \pi, sowie\ 0 \le p(x) \le 1$$

Die Summe aller Wahrscheinlichkeiten – hier, da es sich um eine stetige Funktion handelt, das Integral unter der Kurve der Funktion – muss eins sein. Salopp gesagt muss ein Ereignis x „auf jeden Fall" eintreffen.

Wir können also das Kurvenintegral von $-\infty$ bis ∞ auf eins setzen. Da der Bereich von x begrenzt ist, können die Grenzen von x direkt für das Integral genutzt werden:

$$\int_0^\pi p(x)dx = 1$$

$$\leftrightarrow c\,[-\cos(x)]_0^\pi = c(-\cos(\pi) + \cos(0)) = 2c = 1 \quad \rightarrow c = \frac{1}{2}$$

b) Expected Value:

$$< X >_p = \int p(x)x\,dx = \int_0^\pi \frac{1}{2}\sin(x)\,x\,dx = \left[\frac{1}{2}\sin(x) - x\cos(x)\right]_0^\pi$$

$$= \frac{1}{2}[\sin(\pi) - \pi\cos(\pi) - \sin(0) + 0\cos(o)]_0^\pi = \frac{\pi}{2}$$

c) Variance:

$$Var(X) = \langle X_p^2 \rangle - \langle X_p \rangle^2$$

$$\langle X_p^2 \rangle = \int p(x)x^2 dx = \int \frac{1}{2}\sin(x)\,x^2 dx = \frac{1}{2}[2x\sin(x) + (2-x^2)\cos(x)]_0^\pi = \frac{\pi^2}{2}$$

$$\langle X_p \rangle^2 = \frac{\pi^2}{4}$$

$$Var(x) = \frac{\pi^2}{2} - \frac{\pi^2}{4} = \frac{\pi^2}{4}$$

1.2 Marginal densities

$$p_x(x) = \int_{-\infty}^{\infty} p_{X,Y}\,(x, y')dy'$$
$$= \int_0^1 \frac{3}{7}(2-x)(x+y')dy' = \frac{3}{7}\left[2xy' + y'^2 - x^2y' - \frac{1}{2}xy'^2\right]_0^1$$
$$= \frac{3}{7}\left(-x^2 + \frac{3}{2}x + 1\right) = \frac{3}{7}(2-x)(\frac{1}{2}+x)$$

$$p_y(y) = \int_{-\infty}^{\infty} p_{X,Y}\,(x', y)dx'$$
$$= \int_0^2 \frac{3}{7}(2x' + 2y - x'^2 - x'y)dx' = \frac{3}{7}\left[x'^2 + 2yx' - \frac{1}{3}x'^3 - \frac{1}{2}\,x'^2y\right]_0^2$$
$$= \frac{3}{7}\left(2y + \frac{4}{3}\right)$$

The two variables are independent if and only if p(x,y) = p(x) * p(y).

$$p(x) * p(y) = \frac{3}{7}(2-x)\left(\frac{1}{2}+x\right)\frac{3}{7}\left(2y + \frac{4}{3}\right) = \frac{3}{7}(2-x)(x+y)$$
$$\frac{3}{7}\left(\frac{1}{2}+x\right)\left(2y + \frac{4}{3}\right) = x + y$$

$$\frac{3}{7}y + \frac{2}{7} + \frac{6}{7}xy + \frac{4}{7}x \neq x + y$$

Die Variablen sind somit nicht unabhängig.

1.3 Taylor expansion

$$f(x) = \sqrt{1+x}$$
$$f'(x) = \frac{1}{2}(1+x)^{-\frac{1}{2}}$$
$$f''(x) = -\frac{1}{4}(1+x)^{-\frac{3}{2}}$$
$$f'''(x) = \frac{3}{8}(1+x)^{-\frac{5}{2}}$$
$$T_{3,x_0}(x) = \frac{f(x_0)}{0!}(x-x_0)^0 + \frac{f'(x_0)}{1!}(x-x_0)^1 + \frac{f''(x_0)}{2!}(x-x_0)^2 + \frac{f'''(x_0)}{3!}(x-x_0)^3$$
$$x_0 = 0$$
$$T_{3,0}(x) = \sqrt{1} + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{3}{48}x^3$$

Malte Schwarzer (340819)
Nils Barge (330849)
Bianca Lüders (327210)

1.4 Determinant of a matrix

$$A = \begin{bmatrix} 5 & 8 & 16 \\ 4 & 1 & 8 \\ -4 & -4 & -11 \end{bmatrix}$$

$$\det(A) = 5 * 1 * (-11) + 8 * 8 * (-4) + 16 * 4 * (-4) — 4 * 1 * 16 — 4 * 8 * 5 — 11 * 4 * 8 = 9$$

$$\mathrm{Tr}(A) = 5 + 1 - 11 = -5$$

1.5 Critical points

$$f(x,y) = c + x^2 + y^2$$

$$g(x,y) = c + x^2 - y^2$$

$$\mathrm{grad}(f) = \begin{pmatrix} 2x \\ 2y \end{pmatrix}$$

$$\mathrm{grad}(g) = \begin{pmatrix} 2x \\ -2y \end{pmatrix}$$

$\mathrm{a} = (0,0) => 2 * 0 = 0;\ -2 * 0 = 0$ | a is a critical point

$$H_{f(x,y)} = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

$$H_{g(x,y)} = \begin{pmatrix} g_{xx} & g_{xy} \\ g_{yx} & g_{yy} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$

$$\det(H_f - \lambda I_2) = \det\left(\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}\right)$$

$$= \det\begin{pmatrix} 2-\lambda & 0 \\ 0 & 2-\lambda \end{pmatrix} = (2-\lambda) * (2-\lambda) = \lambda^2 - 4\lambda + 4$$

$\lambda_{1,2} = -\frac{-4}{2} \pm \sqrt{\left(\frac{-4}{2}\right)^2 - 4} = 2 \pm 0 = 2 > 0$ | H_f is positive definite (minimum)

$$\det(\mathrm{H_g} - \lambda I_2) = \det\begin{pmatrix} 2-\lambda & 0 \\ 0 & -2-\lambda \end{pmatrix} = (2-\lambda)(-2-\lambda) = 0$$

$\lambda_1 = 2$

$\lambda_2 = -2$ | H_g is indefinite (no extremum)

1.6 Bayes rule (A=D; +=B)

$$P(D) = 0.01$$
$$P(\overline{D}) = 0.99$$
$$P(+|D) = 0.95$$
$$P(-|\overline{D}) = 0.999$$

$$P(+|\overline{D}) = 0.001$$
$$P(-|D) = 0.05$$

$$P(+|D) * P(D) = P(+ \cap D) = 0.95 * 0.01 = 0.0095$$
$$P\left(+|\overline{D}\right) * P\left(\overline{D}\right) = P\left(+ \cap \overline{D}\right) = 0.001 * 0.99 = 0.00099$$
$$P(+) = P(+ \cap D) + P\left(+ \cap \overline{D}\right) = 0.0095 + 0.00099 = 0.01049$$
$$P(-) = 1 - P(+) = 0.98951$$

$$P(D|+) = \frac{P(+|D) * P(D)}{P(+)} = \frac{0.95 * 0.01}{0.01049} = 0.9056$$
$$P(D|-) = \frac{P(-|D) * P(D)}{P(-)} = \frac{0.05 * 0.01}{0.98951} = 0.0005$$
$$P(\overline{D}|+) = 1 - P(D|+) = 0.0944$$
$$P(\overline{D}|-) = 1 - P(D|-) = 0.9995$$

1.7 Learning paradigms

a)

Supervised learning: predict an outcome based on training data (previous outcomes) – given both the observations and the targets. Allows to use performance measures and to improve classification/regression performance.

Unsupervised learning: find structure or detect pattern solely based on the data itself – no targets given, data-based approach

Reinforcement learning: try something, observe its effect on your goal, adjust and try again – provide the data and a reinforcement signal

b)

To teach a dog to catch a ball (**Reinforced learning**): observation – flight route of the ball, reward – e.g. food

To read hand written addresses from letters (**Supervised learning**): observation – hand written letters, target - actual letter (as it was intended when writing)

To identify groups of users with the same taste of music (**Unsupervised learning**): observations – users, each of them having a unique behavior, purchasing history in the music store, features of the songs they like to listen to