Ex 7-1

a) $$\ln P_W(x_1,\cdots,x_N) = \left[\ln(2\pi\sigma^2)^{-\frac{N}{2}}\right] + \left[-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(x_i-\mu)^2\right]$$

$$\frac{\partial \ln P_W(x_1,\cdots,x_N)}{\partial\mu} = 0 + \left[(-1)\left(-\frac{1}{2\sigma^2}\right)\cdot\sum_{i=1}^{N}2(x_i-\mu)\right]$$

$$= \frac{1}{\sigma^2}\sum_{i=1}^{N}(x_i-\mu) = \frac{1}{\sigma^2}\left[\left(\sum_{i=1}^{N}x_i\right)-N\cdot\mu\right] = 0$$

$$\Rightarrow \hat{\mu}_{ML} = \frac{1}{N}\sum_{i=1}^{N}x_i = \bar{x}.$$

$$\frac{\partial \ln P_W(x_1,\cdots,x_N)}{\partial\sigma} = \left(-\frac{N}{2}\right)\left(\frac{1}{2\pi\sigma^2}\right)(4\pi\sigma) + \left[\frac{1}{4\sigma^4}\cdot 4\sigma\cdot\sum_{i=1}^{N}(x_i-\mu)^2\right]$$

$$= -\frac{N}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{N}(x_i-\mu)^2 = 0$$

$$\Rightarrow \hat{\sigma}^2_{ML} = \frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2 = Cov(x).$$

b) check notebook.

Ex 7-2

a) $P(X=x_i) \Rightarrow$
$P(X=1) = 0.1+0.15+0.25 = 0.5$.
$P(x=2) = 0.05+0.3+0.15 = 0.5$.

$P(Y=y_i) \Rightarrow$
$P(Y=1) = 0.1+0.05 = 0.15$.
$P(Y=2) = 0.15+0.3 = 0.45$.
$P(Y=3) = 0.25+0.15 = 0.4$.

b) $E(X) = \sum_{x_i} x_i \cdot P(X=x_i) = 1 \cdot P(X=1) + 2 \cdot P(X=2)$

$= 0.5 + 2 \times 0.5 = 1.5$.

$E(Y) = \sum_{y_i} y_i \cdot P(Y=y_i) = 1 \times 0.15 + 2 \times 0.45 + 3 \times 0.4$

$= 0.15 + 0.9 + 1.2 = 2.25$.

c) $\mathrm{var}(X) = \sum_{x_i} \cancel{[x_i - E(X)]^2 P(X=x_i) (1-1.5)^2 + (2-1.5)^2}$

$\cancel{= 0.25 + 0.25 = 0}$

$= \sum_{x_i} [x_i - E(x_i)]^2 \cdot P(X=x_i) = (1-1.5)^2 \times 0.5 + (2-1.5)^2 \times 0.5$

$= (0.25 + 0.25) \times 0.5$

$= 0.25$.

$\mathrm{var}(Y) = \sum_{y_i} [y_i - E(Y)]^2 P(Y=y_i)$

$= (1-2.25)^2 \times 0.15 + (2-2.25)^2 \times 0.45$
$+ (3-2.25)^2 \times 0.4$

$= 1.25^2 \times 0.15 + (0.25)^2 \times 0.45 + 0.75^2 \times 0.4$

$= 1.5625 \times 0.15 + 0.0625 \times 0.45 + 0.5625 \times 0.4$

$= 0.4875$.

1.25 × 1.25: 625, 250, 125, = 1.5625

0.75 × 0.75: 375, 525, = 5625

$0.5^4 =$

25 × 25: 125, 50, = 625

$\mathrm{Cov}(X,Y) = E\{[X-E(X)][Y-E(Y)]\}$.

$= \sum_{x_i} \sum_{y_i} [(x_i - E(X))(y_i - E(Y))\, P(X=x_i, Y=y_i)]$

$= 0.1 \times (1-1.5)(1-2.25) + 0.15(1-1.5)(2-2.25)$
$+ \dots + 0.15 \times (2-1.5)(3-2.25)$

$= -0.025$.

BRUNNEN

d) $\rho = \frac{cov(X,Y)}{\sqrt{Var(X)\,var(Y)}} = \frac{-0.025}{\sqrt{0.25 \cdot 0.4875}} \approx -0.07$

e) No, ~~…~~. $p(X)\,p(Y) \neq p(X,Y)$

Ex 7-3

a) – detect hidden correlations by transformation $x \in \mathbb{R}^M \to U^t x \in \mathbb{R}^r$.
in order to reduce complexity and computational cost compared to the original space.

– remove redundant and noisy features.

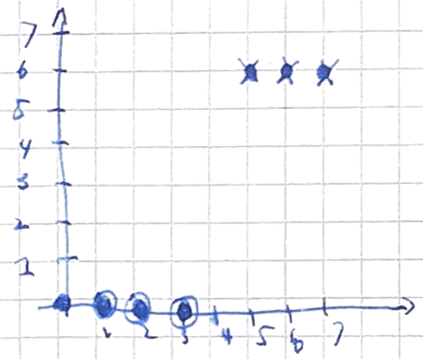
– interpretation and visualization.

– data compression.

When ~~why~~ PCA helpful? Linear correlation. otherwise LDA, RCA, ...

b) N/A.

Ex 7-4



① compute μ.

$\cancel{\mu = X - \bar{X}}$. $\mu = \begin{pmatrix} \bar{X} \\ \bar{X} \\ \vdots \end{pmatrix}$

② $\hat{\Sigma}$. $\tilde{X} = X - \mu$

$Cov(X) \approx \hat{\Sigma} = \frac{1}{N} \tilde{X}^T \tilde{X}$.

③ Eigenvalues/vectors.

④ U = eigenvectors.

$\hat{X} = \mu + \tilde{X} U U^T$. $Y = U^T \tilde{X}$

BRUNNEN