Homework 3:

Problem 1:

Univariate Gaussian PDF : $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ \frac{-(x-\mu)^2}{2\sigma^2} \right\}$

a) prove $f(x)$ is normalized ~ prove $\int f(x)dx = 1$

$$\int f(x)dx = \frac{1}{\sqrt{2\pi\sigma^2}} \int \exp\left\{ \frac{2(x-\mu)^2}{\sigma^2} \right\} dx$$

let $z = \frac{x-\mu}{\sigma} \rightarrow dz = \frac{1}{\sigma} dx$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \int \exp\left\{ \frac{-z^2}{2} \right\} \sigma . dz$$

let $u = \frac{z}{\sqrt{2}} \rightarrow du = \frac{1}{\sqrt{2}} . dz$

$$= \frac{1}{\sqrt{2\pi}} \int \exp(-u^2) . \sqrt{2} . du$$

$$= \frac{1}{\sqrt{\pi}} . \sqrt{\pi}$$

$$= 1.$$

Gaussian Intergral

let $I = \int \exp(-x^2)\, dx$

$\rightarrow I^2 = \iint \exp(-x^2) dx dy$

$dx.dy = r.dr.d\theta$

$\rightarrow I^2 = \int_0^{2\pi}\int_0^{\infty} \exp(-r^2) r.dr.d\theta$

$= 2\pi \int_0 \exp(-r^2)\, r\, dr$

$= 2\pi \int \exp(-x) \frac{1}{2} dx$ $\quad x = r^2$

$= \pi \left[-\exp(-x) \right]_0^{\infty}$

$= \pi$

$\rightarrow I^2 = \pi \rightarrow I = \sqrt{\pi}$

b) prove $E(X) = \mu$.

$$E(X) = \int x\, f(x). dx$$

let $u = \frac{x-\mu}{\sqrt{2\sigma^2}} \rightarrow du = \frac{1}{\sqrt{2\sigma^2}} . dx$

$$= \int x . \frac{1}{\sqrt{2\pi\sigma^2}} . \exp\left(\frac{-(x-\mu)^2}{2\sigma^2} \right) dx$$

$\rightarrow x = \sqrt{2\sigma^2} . u + \mu$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \int \left(\sqrt{2\sigma^2} . u + \mu \right) \exp(-u^2) . \sqrt{2\sigma^2} . du$$

$$= \frac{\sqrt{2\sigma^2}}{\sqrt{\pi}} \int_{-\infty}^{\infty} u . \exp(-u^2)\, du + \frac{\mu}{\sqrt{\pi}} \int \exp(-u^2)\, du$$

let $t = -u^2 \rightarrow dt = -2u . du$

$$= \frac{\sqrt{2\sigma^2}}{\sqrt{\pi}} . \frac{1}{2} \int \exp(t) . dt + \frac{\mu}{\sqrt{\pi}} . \sqrt{\pi}$$

$$= \frac{\sqrt{2\sigma^2}}{\sqrt{\pi}} . \frac{1}{2} \exp(-u^2) \Big|_{-\infty}^{\infty} + \mu = 0 + \mu = \mu$$

c) prove $sd(X) = \sigma^2 \sim$ prove $V(X) = \sigma^2$

$$V(X) = \int (x-\mu)^2 f(x)\,dx$$

$$= \int (x-\mu)^2 \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \cdot dx \qquad \text{let } u = \frac{x-\mu}{\sqrt{2\sigma^2}} \rightarrow du = \frac{1}{\sqrt{2\sigma^2}} \cdot dx$$

$$\rightarrow x-\mu = \sqrt{2\sigma^2} \cdot u$$

$$\rightarrow (x-\mu)^2 = 2\sigma^2 \cdot u^2$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \int 2\sigma^2 \cdot u^2 \cdot \exp(-u^2)\sqrt{2\sigma^2} \cdot du$$

$$= \frac{2\sigma^2}{\sqrt{\pi}} \int u^2 \cdot \exp(-u^2)\,du$$

$$\begin{cases} u = u \\ dv = u \cdot \exp(-u^2) \cdot du \end{cases} \qquad \begin{cases} du = du \\ v = -\frac{1}{2}\exp(-u^2) \end{cases}$$ (proven above)

$$= \frac{2\sigma^2}{\sqrt{\pi}} \left[\left. \frac{-1}{2} \cdot u \cdot \exp(-u^2) \right|_{-\infty}^{\infty} + \frac{1}{2} \int \exp(-u^2) \cdot du \right]$$

$$= \frac{2\sigma^2}{\sqrt{\pi}} \left(0 + \frac{\sqrt{\pi}}{2}\right) = \sigma^2$$

Problem 2:

Multi-variate Gaussian PDF : $f(x) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\left(\frac{-1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right)$

a) prove $f(x)$ is normalized $\sim$ prove $\int_{\mathbb{R}^n} f(x) \cdot dx = 1$.

$$\int f(x) \cdot dx = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \int \exp\left(\frac{-1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right) \cdot dx \qquad \text{let } y = x-\mu \rightarrow dy = dx$$

$$= \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \cdot \frac{1}{2} \int \exp\left(-y^T \Sigma^{-1} y^T\right) dy$$

$$= \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \cdot \int \exp\left[\frac{-z_1^2}{2\lambda_1} + \dots + \frac{-z_n^2}{2\lambda_n}\right] dz$$

(1)

$$= \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \cdot \int \prod \exp\left(\frac{-z_i^2}{2\lambda_i}\right) dz$$

$$= \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \cdot \prod_{i=0}^{n} \sqrt{2\pi\lambda_i}$$

$$= \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \cdot \prod^{n} \sqrt{2\pi} \cdot \prod^{n} \sqrt{\lambda_i}$$

$$= \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \cdot \sqrt{(2\pi)^n} \cdot \sqrt{|\Sigma|}$$

(2)

$$= 1.$$

(1) $y^T\Sigma^{-1}y = y^T U\Lambda^{-1}U^T y$ (3)

$$= (U^Ty)^T\Lambda^{-1}(U^Ty) = Z\Lambda^{-1}Z^T$$

$$= \begin{bmatrix} z_1 \dots & z_n\end{bmatrix} \begin{bmatrix} \frac{1}{\lambda_1} & & 0 \\ \vdots & \ddots & \vdots \\ 0 & & \frac{1}{\lambda_n}\end{bmatrix}\begin{bmatrix} z_1 \\ \vdots \\ z_n\end{bmatrix}$$

$$= \frac{z_1^2}{\lambda_1} + \frac{z_2^2}{\lambda_2} + \dots + \frac{z_n^2}{\lambda_n}$$

$$E = -\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)$$

$$= -\frac{1}{2}\left[x^TAx - 2x^TA\mu + \text{const}\right]$$

$$= -\frac{1}{2}x^TAx + x^TA\mu + \text{const}$$

(2) $|\Sigma| = |U\Lambda U^T| = |U|\cdot|\Lambda|\cdot|U^T| = |U^TU|\cdot|\Lambda| = |I|\cdot|\lambda| = \prod^{n}\lambda_i$

(3) Σ is sym & positive definite $\rightarrow$ $\Sigma = U\Lambda U^T$ (SVD) where U is orthogonal $\rightarrow U^TU = I \rightarrow |U^TU| = 1$

Λ is diagonal matrix of eigenvalues $\lambda_i \rightarrow |\Sigma| = \prod_{i=1}^{n}\lambda_i$

$\downarrow$

Σ^{-1} $\rightarrow$ $\Sigma^{-1} = U\Lambda^{-1}U^T$ where U ...

Λ^{-1} .... $\frac{1}{\lambda_i}$

b) Find marginal distribution.

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \qquad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \rightarrow \Sigma^{-1} = A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

$$g(x_1,x_2) = -\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)$$

$$= -\frac{1}{2}(x_1-\mu_1)^TA_{11}(x_1-\mu_1) - \frac{1}{2}(x_1-\mu_1)^TA_{12}(x_2-\mu_2)$$

$$-\frac{1}{2}(x_2-\mu_2)^TA_{21}(x_1-\mu_1) - \frac{1}{2}(x_2-\mu_2)^TA_{22}(x_2-\mu_2)$$

$$= -\frac{1}{2}\Big[\; x_1^TA_{11}x_1 - x_1^TA_{11}\mu_1 - \mu_1^TA_{11}x_1 + \mu_1^TA_{11}\mu_1$$
$$+\; x_1^TA_{12}x_2 - x_1^TA_{12}\mu_2 - \mu_1^TA_{12}x_2 + \mu_1^TA_{12}\mu_2$$
$$+\; x_2^TA_{12}x_1 - x_2^TA_{12}\mu_1 - \mu_2^TA_{12}x_1 + \mu_2^TA_{12}\mu_1$$
$$+\; x_2^TA_{22}x_2 - x_2^TA_{22}\mu_2 - \mu_2^TA_{22}\mu_2 + \mu_2^TA_{22}\mu_2 \Big]$$

const

$$= -\frac{1}{2}\Big[\; x_1^TA_{11}x_1 + 2x_1^TA_{12}x_2 + x_2^TA_{22}x_2$$
$$-\; x_1^TA_{11}\mu_1 - x_1^TA_{12}\mu_2 - x_2^TA_{12}\mu_1 - x_2^TA_{22}\mu_2$$
$$-\; \mu_1^TA_{11}x_1 - \mu_1^TA_{12}x_2 - \mu_2^TA_{12}x_1 - \mu_2^TA_{22}\mu_2 \Big] + \text{const}$$

$$= -\frac{1}{2}(x_1^TA_{11}x_1) - x_1^TA_{12}x_2 - \frac{1}{2}x_2^TA_{22}x_2 + x_1^TA_{11}\mu_1 + x_1^TA_{12}\mu_2 + x_2^TA_{12}\mu_1 + x_2^TA_{22}\mu_2 + \text{const}$$

$$= -\frac{1}{2}x_1^TA_{11}x_1 - x_1^TA_{12}x_2 + x_1^TA_{11}\mu_1 + x_1^TA_{12}\mu_2 + x_2^TA_{12}\mu_1 + \frac{1}{2}x_2^TA_{22}\mu_2 + \text{const}$$

$$= -\frac{1}{2}x_1^TA_{11}x_1 + x_1^T\left[-A_{12}x_2 + A_{11}\mu_1 + A_{12}\mu_2\right] + \text{const}$$

$$= -\frac{1}{2}x_1^TA_{11}x_1 + x_1^T\left[A_{11}\mu_1 - A_{12}(x_2-\mu_2)\right] + \text{const} \sim -\frac{1}{2}x^TAx + x^TA\mu + \text{const}$$

$\rightarrow x_1 \sim$ Gaussian

Schur complement (proof: (4))

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} M & -MBD^{-1} \\ -D^{-1}CMD^{-1} & D^{-1}CMBD^{-1} \end{bmatrix}, \quad M = (A - BD^{-1}C)^{-1}$$

$$\rightarrow \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}^{-1} = \begin{bmatrix} (\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})^{-1} & -(\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})^{-1}\Sigma_{12}\Sigma_{22}^{-1} \\ \sim & \Sigma_{22}^{-1}\Sigma_{21}(\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})^{-1}\Sigma_{12}\Sigma_{22}^{-1} \end{bmatrix}$$

$$p(x_2) = \int f(x_1, x_2)\, dx_1 = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \int \exp(g(x_1, x_2))\, dx_1$$

$$= \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \int \exp\left[-\frac{1}{2} x_1^T A_{11} x_1 + x_1^T [A_{11}\mu_1 - A_{12}(x_2 - \mu_2)] + const \right] dx_1$$

$$= \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \int \exp\left[\frac{1}{2} (x_1 - A_{11}^{-1} m)^T A_{11} (x_1 - A_{11}^{-1} m) \right] dx_2 \qquad (5)$$

$$= \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \int \exp\Big[\frac{1}{2} [A_{11}\mu_1 - A(x_2 - \mu_2)]^T A_{11}^{-1} [A_{11}\mu_1 - A_{12}(x_2 - \mu_2))$$
$$- \frac{1}{2} x_2^T A_{22} x_2 + x_2^T (A_{22}\mu_2 + A_{12}\mu_1) + const \Big] \qquad (6)$$

$$= \text{————} \Big[-\frac{1}{2} x_2^T (A_{22} - A_{21} A_{11}^{-1} A_{12}) x_2$$
$$+ x_2^T (A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1} \mu_2 + const$$

$$\rightarrow cov(x_2) = \Sigma_2 = (A_{22} - A_{22} \cdot A_{21}^{-1} A_{21})^{-1} = \Sigma_{22}$$
$$E(x_2) = \mu_2 = \Sigma_2 (A_{22} - A_{22} \cdot A_{21}^{-1} A_{21}) \mu_a = \mu_a$$

$$(4) \quad \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = I = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} M & -MBD^{-1} \\ -D^{-1}CM & D^{-1} + D^{-1}CMBD^{-1} \end{pmatrix}$$

top left: $AM + B(-D^{-1}CM) = (A - BD^{-1}C)M = M^{-1}M = I$

top right: $A(-MBD^{-1}) + B(D^{-1} + D^{-1}CMBD^{-1})$
$= -AMBD^{-1} + BD^{-1} + BD^{-1}(CMBD^{-1})$
$= (-AM + BD^{-1}CM)BD^{-1} + BD^{-1}$
$= -IBD^{-1} + BD^{-1}$
$= 0$

$$\begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

bottom left: $CM + D(-D^{-1}CM) = CM - CM = 0$

bottom right: $C(-MBD^{-1}) + D(D^{-1} + D^{-1}CMBD^{-1})$
$= -CMBD^{-1} + I + CMBD^{-1}$
$= I$

(5) $m = A_{11}\mu_1 - A_{12}(x_2-\mu_2) \rightarrow A_{11}^{-1} m = A_{11}^{-1}\left(A_{11}\mu_1 - A_{12}(x_2-\mu_2)\right) = \mu_1 - A_{11}^{-1}A_{12}(x_2-\mu_2)$

$$-\frac{1}{2} x_1^T A_{11} x_1 + x_1^T\left[A_{11}\mu_1 - A_{12}(x_2-\mu_2)\right] + \text{const} = -\frac{1}{2} x_1^T A_{11} x_1 + x_1^T m$$

$$= -\frac{1}{2}(x_1 - A_{11}^{-1} m)^T A_{11} (x_1 - A_{11}^{-1} m) + \frac{1}{2} m^T A_{11}^{-1} m$$

(6) $-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}) =$

$$-\frac{1}{2}(\mathbf{x}_a-\boldsymbol{\mu}_a)^{\mathrm{T}}\boldsymbol{\Lambda}_{aa}(\mathbf{x}_a-\boldsymbol{\mu}_a) - \frac{1}{2}(\mathbf{x}_a-\boldsymbol{\mu}_a)^{\mathrm{T}}\boldsymbol{\Lambda}_{ab}(\mathbf{x}_b-\boldsymbol{\mu}_b)$$
$$-\frac{1}{2}(\mathbf{x}_b-\boldsymbol{\mu}_b)^{\mathrm{T}}\boldsymbol{\Lambda}_{ba}(\mathbf{x}_a-\boldsymbol{\mu}_a) - \frac{1}{2}(\mathbf{x}_b-\boldsymbol{\mu}_b)^{\mathrm{T}}\boldsymbol{\Lambda}_{bb}(\mathbf{x}_b-\boldsymbol{\mu}_b).$$

c) Find conditional distribution.

$$p(x_1|x_2) = \frac{p(x_1,x_2)}{p(x_2)} = \frac{\mathcal{N}(x;\mu,\Sigma)}{\mathcal{N}(x_2;\mu_2,\Sigma_{22})}$$

$$= \frac{1/\sqrt{(2\pi)^n|\Sigma|}\cdot \exp\left[-\frac{1}{2}(x-\mu)\Sigma^{-1}(x-\mu)\right]}{1\sqrt{(2\pi)^{n_2}|\Sigma_{22}|}\exp\left[-\frac{1}{2}(x_2-\mu_2)\Sigma_{22}^{-1}(x_2-\mu_2)\right]}$$

$$= \frac{1}{\sqrt{(2\pi)^{n-n_2}}}\sqrt{\frac{|\Sigma_{22}|}{|\Sigma|}}\exp\left[-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu) + \frac{1}{2}(x_2-\mu_2)^T\Sigma_{22}^{-1}(x_2-\mu_2)\right]$$

$$= \frac{1}{\sqrt{(2\pi)^{n-n_2}}}\sqrt{\frac{|\Sigma_{22}|}{|\Sigma|}}\exp\left\{-\frac{1}{2}\Big[(x_1-\mu_1)^T A_{11}(x_1-\mu_1) + 2(x_1-\mu_1)^T A_{12}(x_2-\mu_2) + (x_2-\mu_2)^T A_{22}(x_2-\mu_2)\Big]\right\}$$

$$= \frac{1}{\sqrt{(2\pi)^{n-n_2}}}\sqrt{\frac{|\Sigma_{22}|}{|\Sigma|}}\exp\left\{-\frac{1}{2}\Big[(x_1-\mu_1)^T(\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})^{-1}(x_1-\mu_1) + 2(x_1-\mu_1)^T - (\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})^{-1}\Sigma_{12}\Sigma_{22}^{-1}(x_2-\mu_2) + (x_2-\mu_2)^T\Sigma_{22}^{-1}\Sigma_{21}(\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})^{-1}\Sigma_{12}\Sigma_{22}^{-1}(x_2-\mu_2)\Big]\right\}$$

$$\overset{(7)}{=} \frac{1}{\sqrt{(2\pi)^{n_1}|\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}|}}$$
$$\exp\left\{-\frac{1}{2}\Big[\big[(x_1-\mu_1)-\Sigma_{12}\Sigma_{22}(x_2-\mu_2)\big]^T(\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})^{-1}\big[\Sigma_{12}\Sigma_{22}^{-1}(x_2-\mu_2)\big]\Big]\right\}$$

$$\rightarrow \quad E(x_1|x_2) = \mu_{x_1|x_2} = \mu_1 + \Sigma_{12}\Sigma_{22}(x_2-\mu_2)$$
$$\Sigma_{x_1|x_2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$$

(7) $\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |D|\cdot|A - BD^{-1}C| \rightarrow \begin{vmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{vmatrix} = |\Sigma_{22}|\cdot|\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}|$