Chapter 1

Wednesday, January 23, 2019 4:31 PM

Problem 1

A) $x_1, \dots, x_n \sim \text{bernulli}(\omega)$, Suppose $\omega \sim \text{Beta}(a,b)$

$$p(\omega) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \omega^{a-1}(1-\omega)^{b-1}$$

Derive the posterior $p(\omega \mid x_{1:n})$

$f(\omega) \propto g(\omega)$

$f(\omega) = c\, g(\omega)$ where c is a constant

$p(\omega \mid x_{1:n}) \propto p(\omega) p(x_{1:n} \mid \omega)$ For Bayes rule, $c = \frac{1}{p(x_{1:n})} = \frac{1}{\int p(x \mid \omega) p(\omega) d\omega}$

$$p(x_{1:n} \mid \omega) = \prod_{i=1}^{n} \omega^{x_i}(1-\omega)^{1-x_i}$$

$$p(\omega \mid x_{1:n}) \propto \omega^{a-1}(1-\omega)^{b-1}\omega^{s}(1-\omega)^{n-s} \qquad s = \sum_{i=1}^{n} x_i$$

$$= \omega^{a+s-1}(1-\omega)^{n-s+b-1}$$

$$\sim \text{beta}(a+s,\ n-s+b)$$

b) The PDF of a gamma RV $x \sim GA(a,b)$ is

$$p(x) = \frac{b^a}{\Gamma(a)} x^{a-1} \exp(-bx)$$

Suppose $x_1 \sim Ga(a_1, 1)$ $x_2 \sim Ga(a_2, 1)$. Define

$$y_1 = \frac{x_1}{x_1 + x_2} \qquad y_2 = x_1 + x_2.$$

Find the joint density for (y_1, y_2) using a direct PDF transformation (& its jacobian). Use this method to find $p(y_1), p(y_2)$ & propose a method for simulating beta RVs assuming you've got gamma RVs.

$x_1 \sim ga(a_1, 1)$ $x_2 \sim ga(a_2, 1)$ To get gamma RVs, add two $\exp(\lambda_1) + \exp(\lambda_2)$

$y_1 = \frac{x_1}{x_1 + x_2}$ $y_2 = x_1 + x_2$ To get exponential, use uniforms w/ inverse logit function

$A: \{(x_1, x_2) : x_1 \in [0, \infty) \times x_2 \in [0, \infty)\}$

$B: \{(y_1, y_2) : y_1 \in [0, 1] \times y_2 \in [0, \infty)\}$

one to one & onto.

$x_1 = h_1(y_1, y_2) = y_1 y_2$ $x_2 = h_2(y_1, y_2) = y_2 - y_1 y_2$

$$J = \begin{vmatrix} \frac{\partial h_1}{\partial y_1} & \frac{\partial h_1}{\partial y_2} \\ \frac{\partial h_2}{\partial y_1} & \frac{\partial h_2}{\partial y_2} \end{vmatrix} = \begin{vmatrix} y_2 & y_1 \\ -y_2 & 1-y_1 \end{vmatrix} = y_2(1-y_1) + y_1 y_2 = y_2 - y_1 y_2 + y_1 y_2$$

$$= y_2$$

$$f_{y_1 y_2}(y_1, y_2) = f_{x_1 x_2}(h_1, h_2)|J|$$

$$f_{x_1 x_2}(x_1, x_2) = \frac{1^{a_1}\, 1^{a_2}}{\Gamma(a_1)\Gamma(a_2)} x_1^{a_1-1}\exp(-x_1)\, x_2^{a_2-1}\exp(-x_2)$$

$$f_{y_1 y_2}(y_1, y_2) = \frac{1}{\Gamma(a_1)\Gamma(a_2)}(y_1 y_2)^{a_1-1}\exp(-y_1 y_2)(y_2 - y_1 y_2)^{a_2-1}\exp(-y_2 + y_1 y_2)(y_2)$$

$$= \frac{1}{\Gamma(a_1)\Gamma(a_2)} y_1^{a_1-1} y_2^{a_1} \exp(\cancel{-y_1 y_2})(y_2)^{a_2-1}(1-y_1)^{a_2-1}\exp(-y_2)\exp(\cancel{y_1 y_2})$$

$$f_{y_1,y_2}(y_1,y_2) = \frac{1}{\Gamma(a_1)\Gamma(a_2)} (y_1y_2)^{a_1-1}\exp(-y_1y_2)(y_2-y_1y_2)^{a_2-1}\exp(-y_2+y_1y_2)(y_2)$$

$$= \frac{1}{\Gamma(a_1)\Gamma(a_2)} y_1^{a_1-1} y_2^{a_1} \cancel{\exp(-y_1y_2)} (y_2)^{a_2-1}(1-y_1)^{a_2-1}\exp(-y_2)\cancel{\exp(y_1y_2)}$$

$$= \frac{1}{\Gamma(a_1)\Gamma(a_2)} y_1^{a_1-1}(1-y_1)^{a_2-1} y_2^{a_1+a_2-1}\exp(-y_2)$$

$$p(y_1) = \frac{1}{\Gamma(a_1)\Gamma(a_2)} \int_0^\infty y_1^{a_1-1}(1-y_1)^{a_2-1} y_2^{a_1+a_2-1}\exp(-y_2)\,dy_2$$

$$= \frac{1}{\Gamma(a_1)\Gamma(a_2)} y_1^{a_1-1}(1-y_1)^{a_2-1} \Gamma(a_1+a_2)$$

$$= \frac{\Gamma(a_1+a_2)}{\Gamma(a_1)\Gamma(a_2)} y_1^{a_1-1}(1-y)^{a_2-1} \sim beta(a_1,a_2)$$

$$p(y_2) = \frac{1}{\Gamma(a_1)\Gamma(a_2)} y_2^{a_1+a_2-1}\exp(-y_2)\int_0^1 y_1^{a_1-1}(1-y_1)^{a_2-1}dy_1$$

$$= \frac{1}{\cancel{\Gamma(a_1)\Gamma(a_2)}} y_2^{a_1+a_2-1}\exp(-y_2)\frac{\cancel{\Gamma(a_1)\Gamma(a_2)}}{\Gamma(a_1+a_2)}$$

$$= \frac{1}{\Gamma(a_1+a_2)} y_2^{a_1+a_2-1}\exp(-y_2) \sim gamma(a_1+a_2, 1)$$

To generate a beta (α,β) transform $ga(\alpha,1)$, $ga(\beta,1)$

C) Suppose that we take independent $x_{1:n}$ from a normal sampling model w/ unknown mean θ and known variance σ^2 $x_i \sim N(\theta,\sigma^2)$. Suppose that θ is given a normal prior dist'n with mean m and variance v. Derive $p(\theta|x_{1:N})$

- normals are nice b/c precisions add together
- posterior mean is convex combination (linear combo where weights add to 1) of data mean and prior mean.

$$x_{1:n} \sim N(\theta,\sigma^2)$$

$$p(\theta) \sim N(m,v) = \frac{1}{\sqrt{2\pi v}}\exp\left(-\frac{1}{2v}(\theta-m)^2\right)$$

$$p(x_{1:n}|\theta,\sigma^2) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{1}{2\sigma^2}(x_i-\theta)^2\right)$$

$$= \frac{1}{(2\pi\sigma)^{n/2}}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\theta)^2\right)$$

$$p(\theta|x_{1:n}) \propto \exp\left(-\frac{1}{2v}(\theta-m)^2 - \frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\theta)^2\right)$$

$$= \exp\left(-\frac{1}{2v}(\theta^2-2\theta m+m^2) - \frac{1}{2\sigma^2}(\sum x_i^2 - 2x_i\theta+\theta^2)\right)$$

$$\propto \exp\left\{-\left(\frac{\theta^2-2\theta m}{2v} + \frac{n\theta^2-2n\bar{x}\theta}{2\sigma^2}\right)\right\}$$

$$= \exp\left\{-\left(\frac{\sigma^2(\theta^2-2\theta m)+v(n\theta^2-2n\bar{x}\theta)}{2v\sigma^2}\right)\right\}$$

$$= \exp\left\{-\frac{1}{2v\sigma^2}(\sigma^2\theta^2-2\sigma^2\theta m+vn\theta^2-2nv\bar{x}\theta)\right\}$$

$$= \exp\left\{-\frac{1}{2v\sigma^2}(\theta^2(\sigma^2+nv)-2\theta(\sigma^2 m - vn\bar{x}))\right\}$$

$$= \exp\left\{-\frac{1}{2v\sigma^2}(\sigma^2+nv)\left(\theta^2-2\theta\left(\frac{\sigma^2 m - v\bar{x}n}{\sigma^2+nv}\right)+\left(\frac{\sigma^2\mu - vn\bar{x}}{\sigma^2+nv}\right)^2 - \left(\frac{\sigma^2\mu - vn\bar{x}}{\sigma^2+nv}\right)^2\right)\right\}$$

$$\propto \exp\left\{-\frac{\sigma^2+nv}{2v\sigma^2}\left(\theta-\left(\frac{\sigma^2 m - vn\bar{x}}{\sigma^2+nv}\right)^2\right)\right\}$$

$$\sim N\left(\frac{\sigma^2 m - vn\bar{x}}{\sigma^2+nv}, \frac{v\sigma^2}{\sigma^2+nv}\right)$$

$$\sim N\left(\frac{\sigma^2 m - v\, n\bar{x}}{\sigma^2 + nv}, \frac{v\sigma^2}{\sigma^2+nv}\right)$$

D) $X_{1:n} \overset{iid}{\sim} N(\theta, \sigma^2)$, $\quad w = 1/\sigma^2$

θ known, σ^2 unknown

$$p(x_i \mid \theta, w) = \left(\frac{w}{2\pi}\right)^{1/2} \exp\left\{-\frac{w}{2}(x_i-\theta)^2\right\}$$

w has a gamma prior w/ hyper parameters a and b. Derive the posterior $p(w \mid x_{1:n})$. Reexpress this as a posterior for σ^2, the variance.

$$p(w) = \frac{b^a}{\Gamma(a)} w^{a-1}\exp(-bw) \qquad p(x_{1:n} \mid w, \theta) = \left(\frac{w}{2\pi}\right)^{n/2}\exp\left\{-\frac{w}{2}\Sigma(x_i-\theta)^2\right\}$$

$$p(w \mid x_{1:n}) \propto \frac{b^a}{\Gamma(a)} w^{a-1}\exp(-bw)\, w^{n/2}\exp\left(-\frac{w}{2}\Sigma(x_i-\theta)^2\right)$$

$$\propto w^{a+\frac{n}{2}-1}\exp\left(-w\left(b+\tfrac{1}{2}\Sigma(x_i-\theta)^2\right)\right)$$

$$p(w \mid x_{1:n}) \sim \text{gamma}\left(a+\frac{n}{2}, b+\frac{1}{2}\Sigma(x_i-\theta)^2\right)$$

Thus $p(\sigma^2 \mid x_{1:n}) \sim \text{inv gamma}\left(a+\frac{n}{2}, b+\frac{1}{2}\Sigma(x_i-\theta)^2\right)$

E) $X_{1:n} \sim N(\theta, \sigma_i^2)$ $\quad \theta \sim N(m, v)$

θ unknown, σ_i^2 known

Derive $p(\theta \mid x_1, \dots, x_N)$. Express the posterior mean in a form that is clearly interpretable as a weighted avg of the observations and the prior mean

$$p(x_{1:n} \mid \theta, \sigma_i^2) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_i^2}}\exp\left(-\frac{1}{2\sigma_i^2}(x_i-\theta)^2\right)$$

$$= \left(\frac{1}{2\pi}\right)^{n/2}\prod_{i=1}^n\left(\frac{1}{\sigma_i}\right)\exp\left(-\frac{1}{2}\sum_{i=1}^n\frac{1}{\sigma_i^2}(x_i-\theta)^2\right)$$

$$\propto \exp\left(-\frac{1}{2}\Sigma\frac{1}{\sigma_i^2}(x_i-\theta)^2\right)$$

$$p(\theta \mid x_{1:n}) \propto \exp\left(-\frac{1}{2v}(\theta-m)^2 + -\frac{1}{2}\sum_{i=1}^n\frac{1}{\sigma_i^2}(x_i^2 - 2x_i\theta + \theta^2)\right)$$

$$= \exp\left\{-1\left(\frac{\theta^2 - 2\theta m + m^2}{2v} + \frac{\Sigma\frac{1}{\sigma_i^2}(x_i^2 - 2x_i\theta + \theta^2)}{2}\right)\right\}$$

$$\propto \exp\left\{-\left(\frac{\theta^2 - 2m\theta + v\Sigma\frac{1}{\sigma_i^2}\theta^2 - v\Sigma\frac{1}{\sigma_i^2}2x_i\theta}{2v}\right)\right\}$$

$$= \exp\left\{-\frac{1}{2v}\left(1+v\Sigma\frac{1}{\sigma_i^2}\right)\left(\theta^2 - 2\theta\left(\frac{m+v\Sigma\frac{1}{\sigma_i^2}x_i}{1+v\Sigma\frac{1}{\sigma_i^2}}\right) + \left(\frac{m+v\Sigma\frac{1}{\sigma_i^2}x_i}{1+v\Sigma\frac{1}{\sigma_i^2}}\right)^2\right.\right.$$

$$\left(\frac{m+v\Sigma\frac{1}{\sigma_i^2}x_i}{\dots}\right)^2$$

$$\ldots \left\{ - \frac{\ldots}{2v} \left(\ldots \left(\frac{\ldots}{1+v\sum\frac{1}{\sigma_i^2}} \right) \left(\frac{\ldots}{1+v\sum\frac{1}{\sigma_i^2}} \right) - \left(\frac{m+v\sum\frac{1}{\sigma_i^2}x_i}{1+v\sum\frac{1}{\sigma_i^2}} \right)^2 \right) \right\}$$

$$\propto \exp\left\{ - \frac{\left(1+v\sum\frac{1}{\sigma_i^2}\right)}{2v} \left(\theta - \left(\frac{m+v\sum\frac{1}{\sigma_i^2}x_i}{1+v\sum\frac{1}{\sigma_i^2}} \right)^2 \right)\right.$$

$$\sim N\left(\frac{m+v\sum\frac{1}{\sigma_i^2}x_i}{1+v\sum\frac{1}{\sigma_i^2}}, \frac{v}{1+v\sum\frac{1}{\sigma_i^2}} \right)$$

F) Suppose $(x|w) \sim N(m, w^{-1})$ where $w \sim Ga\left(\frac{a}{2}, \frac{b}{2}\right)$ prior.
Show that the marginal dist'n of x is student's t w/d degrees of freedom center m, and scale parameter $(b/a)^{1/2}$

A mixture of normals is also pick a random μ and a random σ^2?

$$p(w) = \frac{(b/2)^{a/2}}{\Gamma(a/2)} w^{a/2-1} \exp(-bw/2)$$

$$p(x|w) = \left(\frac{w}{2\pi}\right)^{1/2} \exp\left\{-\frac{w}{2}(x-m)^2\right\}$$

$$p(x,w) = p(x|w)p(w)$$

$$p(x) = \int_0^\infty p(x,w)\,dw$$

$$= \int_0^\infty \frac{(b/2)^{a/2}}{\Gamma(a/2)} w^{a/2-1} \exp(-bw/2) \left(\frac{w}{2\pi}\right)^{1/2} \exp\left\{-\frac{w}{2}(x-m)^2\right\} dw$$

$$= \frac{(b/2)^{a/2}}{\Gamma(a/2)(2\pi)^{1/2}} \underbrace{\int_0^\infty w^{a/2-1+\frac{1}{2}} \exp\left\{-w\left(\frac{b}{2}+\frac{1}{2}(x-m)^2\right)\right\} dw}_{\text{Kernel of } Ga\left(\frac{a+1}{2}, \left(\frac{b}{2}+\frac{1}{2}(x-m)^2\right)\right)}$$

$$= \frac{(b/2)^{a/2}}{\Gamma(a/2)(2\pi)^{1/2}} \frac{\Gamma\left(\frac{a+1}{2}\right)}{\left(\frac{b}{2}+\frac{(x-m)^2}{2}\right)^{\frac{a+1}{2}}}$$

$$= \frac{\Gamma\left(\frac{a+1}{2}\right)}{\Gamma\left(\frac{a}{2}\right)(2\pi)^{1/2}} \frac{\left(\frac{b}{2}\right)^{a/2} \Big/ \left(\frac{b}{2}\right)^{\frac{a+1}{2}}}{\left(\frac{b}{2}+\frac{(x-m)^2}{2}\right)^{a+1/2} \Big/ \left(\frac{b}{2}\right)^{\frac{a+1}{2}}}$$

$$= \frac{\Gamma\left(\frac{a+1}{2}\right)}{\Gamma\left(\frac{a}{2}\right)(2\pi)^{1/2}\left(\frac{b}{2}\right)^{1/2}} \left(1+\frac{\frac{(x-m)^2}{2}}{\frac{b}{2}}\right)^{\frac{a+1}{2}}$$

$$= \frac{\Gamma\left(\frac{a+1}{2}\right)}{\Gamma\left(\frac{a}{2}\right)(b\pi)^{1/2}\left(\frac{\sqrt{a}}{\sqrt{a}}\right)} \left(\frac{a}{a}\left(1+\frac{(x-m)^2}{b}\right)\right)^{\frac{a+1}{2}}$$

$$= \frac{\Gamma\left(\frac{a+1}{2}\right)}{\Gamma\left(\frac{a}{2}\right)\left(\frac{b}{a}\right)^{1/2}\sqrt{\pi a}} \left(\frac{a+\left(\frac{x-m}{b/a}\right)^2}{a}\right)^{\frac{a+1}{2}}$$

The multivariate normal dist'n

A) $Cov(X)$ of a vector valued RV x is defined as the matrix whose (i,j) entry is the covariance between x_i and x_j. In the matrix notation,

$$cov(x) = E\{(x-\mu)(x-\mu)^T\}$$

where μ is the mean vector whose ith component is $E(x_i)$. Prove:

is $E(X_i)$. Prove:

1.) $Cov(x) = E(xx^T) - \mu\mu^T$
2.) $Cov(Ax+b) = A\,cov(x)A^T$

for matrix A and vector b.

1.) $E\{(x-\mu)(x-\mu)^T\}$ $\quad x_{n\times 1} \quad x\mu^T = n\times n$
$\quad \mu_{n\times 1} \quad \mu x^T = n \times n$

$= E\{XX^T - X\mu^T - \mu X^T + \mu\mu^T\}$

$= E(XX^T) - E(X)\mu^T - \mu E(X^T) + \mu\mu^T$

$= E(XX^T) - \mu\mu^T - \cancel{\mu\mu^T} + \cancel{\mu\mu^T}$

$= E(XX^T) - \mu\mu^T$

2.) $Cov(Ax+b) = E\{(AX+b)(AX+b)^T\} - E(AX+b)E(AX+b)^T$

$= E\{AXX^TA^T + AXb^T + bX^TA^T + bb^T\} - (AE(X)+b)(E(X)^TA^T + b^T)$

$= AE(XX^T)A^T + \cancel{AE(X)b^T} + \cancel{bE(X^T)A^T} + \cancel{bb^T}$
$- AE(X)E(X)^TA^T - \cancel{AE(X)b^T} - \cancel{bE(X^T)A^T} - \cancel{bb^T}$

$= AE(XX^T)A^T - A\mu\mu^TA^T$

$= A(E(XX^T) - \mu\mu^T)A^T$

$= A\,Cov(X)A^T$

B) Consider the random vector $Z = (Z_1, \dots, Z_p)^T$, w/each entry having an independent standard normal dist'n. Derive the PDF and MGF of Z, expressed in vector notation. We say that Z has a standard multivariate normal dist'n.

If we have a random vector $Z = (Z_1, \dots, Z_p)^T$ where $Z_i \overset{iid}{\sim} N(0,1)$, then the distribution of Z is the joint density of Z_i.

$$Z_i \overset{iid}{\sim} N(0,1) \Rightarrow p(Z_i) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} Z_i^2\right)$$

$p(Z) = p(Z_1, \dots, Z_p) = p(Z_1)p(Z_2)\cdots p(Z_p)$ by the independence of Z_i

$$= \prod_{i=1}^{p} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} Z_i^2\right)$$

$$= \frac{1}{(2\pi)^{p/2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{p} Z_i^2\right)$$

$$= \frac{1}{(2\pi)^{p/2}} \exp\left(-\frac{1}{2} Z^TZ\right)$$

Since Z_i are independent, the joint mgf $M(z_1, \dots, z_p)$ is

$$M(z_1, \dots, z_p) = E\left[e^{c_1 z_1 + \dots + c_p z_p}\right] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{c_1 z_1 + \dots + c_p z_p} f(z_1, \dots, z_p)\, dz_1 \dots dz_p$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{c_1 z_1} \cdots e^{c_p z_p} f(z_1) \cdots f(z_p)\, dz_1 \cdots dz_p$$

$$= \int_{-\infty}^{\infty} e^{c_1 z_1} f(z_1)\, dz_1 \cdots \int_{-\infty}^{\infty} e^{c_p z_p} f(z_p)\, dz_p$$

$$= M_{z_1}(c_1) \cdots M_{z_p}(c_p)$$

The mgf of any z_i is

$$E(e^{c_1 z_1}) = \int_{-\infty}^{\infty} e^{c_1 z_1} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} z_1^2\right) dz_1$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(c_1 z_1 - \frac{1}{2} z_1^2\right) dz_1$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}(z_1^2 - 2c_1 + c_1^2 - c_1^2\right) dz_1$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \underbrace{\exp\left(-\frac{1}{2}(z_1 - c_1)^2\right)}_{\text{kernel of } N(c_1, 1)} \exp\left(\frac{1}{2} c_1^2\right) dz_1$$

$$= \frac{1}{\sqrt{2\pi}} \exp\left(\frac{1}{2} c_1^2\right) \sqrt{2\pi}$$

$$= \exp\left(\frac{1}{2} c_1^2\right)$$

Therefore, $M_{\underset{\sim}{z}}(\underset{\sim}{c}) = \prod_{i=1}^{p} \exp\left(\frac{1}{2} c_i^2\right)$

$$= \exp\left(\frac{1}{2} c^T c\right) \quad \text{where } c = (c_1, \dots, c_p)$$

c) A vector valued RV $x = (x_1, \dots, x_p)^T$ has a multivariate normal dist iff every linear combo of its components is univariate normal. $\forall$ vectors a not identically, the scalar $z = a^T x$ is normally distributed. From this def, prove that x is multivariate normal, written $x \sim N(\mu, \Sigma)$, iff its moment-generating function is of the form

$$E(\exp\{t^T x\}) = \exp(t^T \mu + t^T \Sigma t / 2)$$

Suppose X is multivariate normal $\sim N(\mu, \Sigma)$. We can build X from a standard multivariate normal $Z = (z_1, \dots, z_p)$ by

$$X = \Sigma^{1/2} Z + \mu$$

$$E(X) = \Sigma^{1/2} E(Z) + \mu = 0 + \mu$$

$$\mathrm{Cov}(X) = \Sigma^{1/2} \mathrm{Cov}(Z) (\Sigma^{1/2})^T$$

$$= \Sigma^{1/2} \mathrm{cov}(Z) \Sigma^{1/2} \quad \text{because } \Sigma^T = \Sigma$$

$$= \Sigma^{1/2} I \Sigma^{1/2} = \Sigma$$

In part b), we showed the mgf of Z is $\exp(-\frac{1}{2}c^Tc)$.

We also have that for a generic formula for the mgf of $aX+b$

$$M_{aX+b}(t) = e^{bt}M_X(at)$$

So $M_{\Sigma^{1/2}Z+\mu}(t) = e^{\mu^T t}\exp(\frac{1}{2}(\Sigma^{1/2}t)^T(\Sigma^{1/2}t))$

$= e^{\mu^T t}\exp(\frac{1}{2}t^T\Sigma^{1/2}\Sigma^{1/2}t)$ $\quad$ $\mu^T t$ is a scalar so $\mu^T t = t^T\mu$

$= e^{t^T\mu}\exp(\frac{1}{2}t^T\Sigma t)$

Now to show the backwards direction, assume X has the mgf

$$M_X(t) = e^{t^T\mu}\exp(\frac{1}{2}t^T\Sigma t)$$

then for any linear combination of the components of X, a^TX, the mgf of a^TX is:

$M_{a^TX}(t) = e^{t^T\mu}\exp(\frac{1}{2}(a^Tt)^T\Sigma(a^Tt))$ $\quad$ $E(a^TX) = a^T\mu$

$= e^{t^T\mu}\exp(\frac{1}{2}t^Ta\Sigma a^Tt)$ $\quad$ $Cov(a^TX) = a^T Cov(X)a$

$= e^{t^T\mu}\exp(\frac{1}{2}t^T tr(a\Sigma a^T)t)$ $\quad$ $= a^T\Sigma a$

$\quad$ $= tr(a^T\Sigma a)$

$= e^{(a^Tt)^T\mu}\exp(\frac{1}{2}t^Tt\, tr(a\Sigma a^T))$

$= e^{(a^Tt)^T\mu}\exp(\frac{1}{2}t^Tt\, tr(a\Sigma a^T))$ is the univariate mgf of a linear combination of normals. Because mgfs uniquely characterize the distribution, a^TX is a normal distribution with mean $a^T\mu$ and covariance $tr(a^T\Sigma a)$.

D) A random vector is multivariate normal iff it is an affine transformation of independent univariate normals. Prove the "if" statement. Let z have a standard multivariate dist'n, define the random vector

$$x = Lz + \mu \qquad L_{p\times p}, \; R(L) = p$$

Prove that x is multivariate normal. Use the moment identities proved above to compute the expected value and covariance matrix of X.

Let z have a multivariate standard normal dist'n. This means $[z_1, \ldots, z_p]$, $z_i \overset{iid}{\sim} N(0,1)$.

Let $x = Lz + \mu$.

Then $X_i = \sum_{j=1}^{p} l_{ij}z_j + \mu_i = l_i^T z + \mu$ where l_i is the ith row of L

Then $X_i = \sum_{j=1}^{p} L_{ij} Z_j + \mu_i = L_i^T Z + \mu$ where L_i is the ith row of L

To prove that X_i is normal, we will once again use the uniqueness of the mgf.

$M_{X_i}(t) = M_{L_{i1}Z_j}(t) \cdots M_{L_{ip}Z_j}(t)$ by the independence of the Z_j

Again, invoke $M_{aX+b}(t) = e^{bt} M_X(at)$. We showed before

$$M_{Z_j}(t) = \exp\left(\tfrac{1}{2}t^2\right)$$

therefore, $M_{L_{ij}Z_j}(t) = \exp(\mu_i t)\exp\left(\tfrac{1}{2}(L_{ij}t)^2\right)$

then $M_{X_i}(t) = \prod_{j=1}^{p} \exp(\mu_i t)\exp\left(\tfrac{1}{2}(L_{ij}t)^2\right)$

$$= \exp(t(n\mu_i))\exp\left(\tfrac{1}{2}t^2 \sum_{j=1}^{p} L_{ij}\right)$$

this is the mgf of a $N(n\mu_i, \sum_{j=1}^{p} L_{ij})$ RV.

Thus, $X_1, \ldots, X_p \sim$ Normal. To show that X is multivariate normal, any linear combination of X_i must also be normal. We can prove this using the same logic as above. For example,

$$\begin{aligned} aX_1 + bX_2 &= a(L_1^T Z + \mu) + b(L_2^T Z + \mu) \\ &= aL_1^T Z + a\mu + bL_2^T Z + b\mu \\ &= (aL_1^T + bL_2^T)Z + (a+b)\mu \end{aligned}$$

which is the same form as $L_i^T Z + \mu$ and so is also normal.

Now to find $E(X)$ and $Cov(X)$

$$E(X) = E(LZ + \mu) = LE(Z) + \mu = \mu$$

$$Cov(X) = Cov(LZ + \mu) = L\,cov(Z)L^T = LIL^T = LL^T$$

E) Prove the only if direction. Suppose that x has a multivariate normal dist'n. Prove that x can be written as an affine transformation of standard normal RVs. Propose an algorithm for simulating MVN RVs with a specified mean and covariance matrix.

Suppose X is $MVN(\mu, \Sigma)$.

Starting with $Z \sim MVN(0, I)$, we have already shown that $LZ + \mu$ is a MVN RV.

Since covariance matrices are positive semidefinite, there exists a square root matrix $\Sigma^{1/2}$ such that

$$\Sigma^{1/2}\Sigma^{1/2} = \Sigma$$

Suppose we have an affine linear combination

$$X = \Sigma^{1/2}Z + \mu$$

We know from part D that X is multivariate normal. To prove $X \sim MVN(\mu, \Sigma)$, we find the expectation and variance of $\Sigma^{1/2}Z + \mu$.

$$E(X) = E(\Sigma^{1/2}Z + \mu) \qquad Cov(X) = Cov(\Sigma^{1/2}Z + \mu)$$

$$= \Sigma^{1/2}E(Z) + \mu \qquad = \Sigma^{1/2}Cov(Z)\Sigma^{1/2}$$

$$= \mu \qquad = \Sigma^{1/2} I \Sigma^{1/2} = \Sigma$$

Thus, any $X \sim MVN(\mu, \Sigma)$ can be written as

$$X = \Sigma^{1/2}Z + \mu$$

In order to build X in this way, given μ and Σ, it is necessary to decompose Σ into $\Sigma^{1/2}$. We can do this via spectral decomposition.

Spectral decomposition states that Σ can be decomposed into

$\Sigma = P\Lambda P^{-1}$ where $P = P^{-1} = P^T$ and Λ is a diagonal matrix with the eigenvalues of Σ on the diagonal. The columns of P are the corresponding eigenvectors of Σ.

Since Σ is positive semidefinite, all of its eigenvalues are non negative. Therefore, $\Lambda^{1/2}$ can be constructed by simply taking the square root of each diagonal term of Λ. Thus, we can set

$$\Sigma^{1/2} = P\Lambda^{1/2}P^{-1}$$

We can check that $\Sigma^{1/2}\Sigma^{1/2} = \Sigma$:

$$\Sigma^{1/2}\Sigma^{1/2} = P\Lambda^{1/2}P^{-1}P\Lambda^{1/2}P^{-1}$$

$$\Sigma^{1/2}\Sigma^{1/2} = P\Lambda^{1/2}P^{-1}P\Lambda^{1/2}P^{-1}$$
$$= P\Lambda^{1/2}\Lambda^{1/2}P^{-1}$$
$$= P\Lambda P^{-1} = \Sigma$$

F) Use this last result, together w/ the PDF of a standard multivariate normal, to show that the PDF of a mvn $x \sim MVN(\mu, \Sigma)$ takes the form

$$p(x) = C\exp\{-Q(x-\mu)/2\}$$

for some constant C and quadratic form $Q(x-\mu)$.

The pdf of $Z \sim MVN(0, I)$ is

$$p(Z) = \frac{1}{|2\pi|^{1/2}}\exp\left(-\frac{1}{2}(Z^T Z)\right)$$

To find the pdf of X, I will use the transformation of variables method:

$$f_X(x) = f_Z(g^{-1}(x))\left|\frac{d}{dx}g^{-1}(x)\right| \quad x \in \mathcal{X}$$

$$g(z) = X = \Sigma^{1/2}Z + \mu \Rightarrow g^{-1}(x) = (\Sigma^{1/2})^{-1}(x-\mu)$$

$$\frac{d}{dx}g^{-1}(x) = \frac{d}{dx}(\Sigma^{1/2})^{-1}(x-\mu)$$
$$= (\Sigma^{1/2})^{-1}$$

Thus, $f_X(x) = \frac{1}{|2\pi|^{1/2}}\exp\left(-\frac{1}{2}\left[(\Sigma^{1/2})^{-1}(x-\mu)\right]^T\left[(\Sigma^{1/2})^{-1}(x-\mu)\right]\right)(\Sigma^{1/2})^{-1}$

somehow $= \frac{1}{|2\pi\Sigma|^{1/2}}\exp\left(-\frac{1}{2}(x-\mu)^T\underbrace{((\Sigma^{1/2})^{-1})^T(\Sigma^{1/2})^{-1}}_{=\Sigma^{-1}}(x-\mu)\right)$

$=\Sigma^{-1}$ because $\Sigma^{1/2}$ is symmetric so $((\Sigma^{1/2})^{-1})^T = (\Sigma^{1/2})^{-1}$

$$= \frac{1}{|2\pi\Sigma|^{1/2}}\exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)$$
$$= C\exp\left(-\frac{1}{2}Q(x-\mu)\right)$$

G) Let $x_1 \sim N(\mu_1, \Sigma_1)$ and $x_2 \sim N(\mu_2, \Sigma_2)$ where x_1 & x_2 are independent of each other. Let

$y = Ax_1 + Bx_2$ for matrices A, B of full column rank and appropriate dimension.

Note x_1 and x_2 need not have the same dimension as long as Ax_1 and Bx_2 do. Use the previous results to characterize the distribution of y.

Note X_1 and X_2 need not have the same dimension as long as AX_1 and BX_2 do. Use the previous results to characterize the distribution of y.

Let $X_1 \sim N(\mu_1, \Sigma_1)$ and $X_2 \sim N(\mu_2, \Sigma_2)$

Let X_1 be $p \times 1$ and X_2 be $q \times 1$

Let A be $n \times p$ and B be $n \times q$

Thus AX_1 is $n \times 1$ and BX_2 is also $n \times 1$.

Construct $Y = AX_1 + BX_2$. The distribution of Y is multivariate normal. To show this, let

$$C = AX_1 \quad \text{and} \quad D = BX_2$$

Because X_1 and X_2 are MVN, any linear combination of their components is normal. Therefore every component of C and D are normal. In fact, C and D are both multivariate normal by result E.

We can write X_1 and X_2

$$X_1 = \Sigma_1^{1/2} Z_1 + \mu_1 \qquad X_2 = \Sigma_2^{1/2} Z_2 + \mu_2$$

$$C = AX_1 = A\Sigma_1^{1/2} Z_1 + A\mu_1 \qquad D = BX_2 = B\Sigma_2^{1/2} Z_2 + B\mu_2$$

$$C \sim MVN(A\mu_1, A\Sigma_1 A^T) \qquad D \sim MVN(B\mu_2, B\Sigma_2 B^T)$$

by result A $\qquad$ by result A

Since X_1 and X_2 were independent, C and D are also independent. We now have that Y is the sum of two independent MVN RVs.

$$Y = C + D$$

To find the distribution of Y, we can use its mgf. Since C and D are independent, the mgf of $C+D$ is the product of $M_C(t)$ and $M_D(t)$.

$$M_C(t) = \exp((A\mu_1)^T t + \tfrac{1}{2} t' A\Sigma_1 A^T t)$$

$$M_D(t) = \exp((B\mu_2)^T t + \tfrac{1}{2} t' B\Sigma_2 B^T t)$$

$$M_Y(t) = M_C(t) M_D(t)$$

$$= \exp((A\mu_1 + B\mu_2)^T t + \tfrac{1}{2} t'(A\Sigma_1 A^T + B\Sigma_2 B^T) t)$$

Thus, $Y \sim MVN(A\mu_1 + B\mu_2, A\Sigma_1 A^T + B\Sigma_2 B^T)$