1. A1.1

let $\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ if all inverses exist

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \qquad XX^{-1} = I$$

$$= \begin{pmatrix} AB_{11} + BB_{21} & AB_{12} + BB_{22} \\ CB_{11} + DB_{21} & CB_{12} + DB_{22} \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}$$

$AB_{12} + BB_{22} = 0$ $\Rightarrow$ $B_{12} = -A^{-1}BB_{22}$ plug in $AB_{11} + BB_{21} = I$ $\Rightarrow$

$CB_{11} + DB_{21} = 0$ $B_{21} = -D^{-1}CB_{11}$ $CB_{12} + DB_{22} = I$

$AB_{11} + B(-D^{-1}CB_{11}) = I$ $\qquad$ $C(-A^{-1}BB_{22}) + DB_{22} = I$

$(A - BD^{-1}C)B_{11} = I$ $\qquad$ $(D - CA^{-1}B)B_{22} = I$

$B_{11} = (A - BD^{-1}C)^{-1}$ $\qquad$ $B_{22} = (D - CA^{-1}B)^{-1}$

Then $\quad B_{12} = -A^{-1}B(D - CA^{-1}B)^{-1}$

$\quad B_{21} = -D^{-1}C(A - BD^{-1}C)^{-1}$

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} (A - BD^{-1}C)^{-1} & -A^{-1}B(D - CA^{-1}B)^{-1} \\ -D^{-1}C(A - BD^{-1}C)^{-1} & (D - CA^{-1}B)^{-1} \end{pmatrix}$$

$$(A - BD^{-1}C)^{-1} = (A(I - A^{-1}BD^{-1}C))^{-1}$$

$$= (I - A^{-1}BD^{-1}C)^{-1}A^{-1}$$

$$= (I - A^{-1}BD^{-1}C)^{-1}(I \underbrace{- A^{-1}BD^{-1}C + A^{-1}BD^{-1}C)}_{\text{cancel}} A^{-1}$$

$$= (I + (I - A^{-1}BD^{-1}C)^{-1}A^{-1}BD^{-1}C)A^{-1}$$

$$= A^{-1} + (I - A^{-1}BD^{-1}C)^{-1}A^{-1}BD^{-1}CA^{-1}$$

matrix identity $X + XYX = X(I + YX) = (I + XY)X$

$$X(I + YX)(I + YX)^{-1} = (I + XY)X(I + YX)^{-1}$$

$$X = (I + XY)X(I + YX)^{-1}$$

$$(I + XY)^{-1}X = X(I + YX)^{-1}$$

treat A^{-1} as X, $BD^{-1}C$ as Y

$$= A^{-1} + A^{-1}(I - BD^{-1}CA^{-1})^{-1}BD^{-1}CA^{-1}$$

treat B as X, $D^{-1}CA^{-1}$ as Y

$$= A^{-1} + A^{-1}B(I - D^{-1}CA^{-1}B)^{-1}D^{-1}CA^{-1}$$

treat D^{-1} as X, $CA^{-1}B$ as Y

$$= A^{-1} + A^{-1}BCD^{-1}C(I - CA^{-1}BD^{-1})^{-1}CA^{-1}$$

$$= A^{-1} + A^{-1}BCD^{-1}C((D - CA^{-1}B)D^{-1})^{-1}CA^{-1}$$

$$= A^{-1} + A^{-1}D^{-1}D(D - CA^{-1}B)^{-1}CA^{-1}$$

$$= A^{-1} + A^{-1}(D - CA^{-1}B)^{-1}CA^{-1}$$

Same for $(D - CA^{-1}B)^{-1} = D^{-1} + D^{-1}C(A - BD^{-1}C)^{-1}BD^{-1}$

Sherman – Morrison

$$A^{-1} - (I + v^TA^{-1}u)^{-1}A^{-1}uv^TA^{-1} = A^{-1} - \frac{A^{-1}uv^TA^{-1}}{I + v^TA^{-1}u}$$ scalar denominator

$$(A + uv^T)(A^{-1} - (I + v^TA^{-1}u)^{-1}A^{-1}uv^TA^{-1})$$

$$= (A + uv^T)\left(A^{-1} - \frac{A^{-1}uv^TA^{-1}}{I + v^TA^{-1}u}\right)$$

$$= I + uv^TA^{-1} - \frac{uv^TA^{-1}uv^TA^{-1} + AA^{-1}uv^TA^{-1}}{1 + v^TA^{-1}u}$$

$$= I + uv^TA^{-1} - \frac{u(1 + v^TA^{-1}u)v^TA^{-1}}{1 + v^TA^{-1}u}$$

$$= I + uv^TA^{-1} - uv^TA^{-1}$$

$$= I$$

similarly $\left(A^{-1} - \frac{A^{-1}uv^TA^{-1}}{I + v^TA^{-1}u}\right)(A + uv^T) = I$

$$\therefore (A + uv^T)^{-1} = A^{-1} - (1 + v^TA^{-1}u)^{-1}A^{-1}uv^TA^{-1}$$

2. A1.2

we have $\frac{\partial a^T x}{\partial x} = a$ $\frac{\partial x^T a}{\partial x} = a$

by chain rule

$$\frac{\partial (f(x,y))}{\partial x} = \frac{\partial (f(x,y))}{\partial x} + \frac{\partial (y(x)^T)}{\partial x} \frac{\partial (f(x,y))}{\partial y}$$

let $y = Ax$

$$\frac{\partial (x^T A x)}{\partial x} = \frac{\partial (x^T y)}{\partial x} + \frac{\partial (y(x)^T)}{\partial x} \frac{\partial (x^T y)}{\partial y}$$

$$= y + \frac{\partial (x^T A^T)}{\partial x} x$$

$$= Ax + A^T x$$

$$= (A + A^T) x$$

3.

A2.1 prove Theorem A2.1

1. let $Z = X + Y$, $X \sim Gamma(a, \theta)$, $Y \sim Gamma(\beta, \theta)$

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\, dx$$

$$= \int_0^{\infty} \frac{\theta^a}{\Gamma(a)} x^{a-1} e^{-\theta x} \frac{\theta^\beta}{\Gamma(\beta)} (z-x)^{\beta-1} e^{-\theta(z-x)}\, dx$$

$$= \frac{\theta^{a+\beta} e^{-\theta z}}{\Gamma(a)\Gamma(\beta)} \int_0^z x^{a-1} \cancel{e^{-\theta x}} (z-x)^{\beta-1} \cancel{e^{\theta x}}\, dx$$

let $w = \frac{x}{z}$ $\quad \frac{dw}{dx} = \frac{1}{z} \Rightarrow dw = z^{-1} dx$ $\quad x=0 \Rightarrow w=0$

$x=z \Rightarrow w=1$

$$= \frac{\theta^{a+\beta} e^{-\theta z}}{\Gamma(a)\Gamma(\beta)} \int_0^z x^{a-1} z^{\beta-1} (1 - \frac{x}{z})^{\beta-1}\, dx$$

$$= \frac{\theta^{a+\beta} e^{-\theta z}}{\Gamma(a)\Gamma(\beta)} \int_0^1 w^{(a-1)} z^{(a-1)} z^{(\beta-1)} (1-w)^{\beta-1} z\, dw$$

$$= \frac{\theta^{a+\beta} e^{-\theta z}}{\Gamma(a)\Gamma(\beta)} \cdot z^{(a-1+\beta-1+1)} B(a,\beta) \cdot \underbrace{\int_0^1 \frac{w^{(a-1)} (1-w)^{\beta-1}}{B(a,\beta)}\, dw}_{\text{density} \Rightarrow 1}$$

$$B(a,\beta) = \frac{\Gamma(a)\Gamma(\beta)}{\Gamma(a+\beta)}$$

$$= \frac{\theta^{\alpha+\beta}}{\Gamma(\alpha+\beta)} z^{\alpha+\beta-1} \cdot e^{-\theta z} \sim \text{Gamma}(\alpha+\beta, \theta)$$

2. let $U = \frac{X}{X+Y}$, $V = X+Y$

$X = UV$, $Y = V - UV$

Jacobian $J = \begin{pmatrix} \frac{\partial X}{\partial U} & \frac{\partial X}{\partial V} \\ \frac{\partial Y}{\partial U} & \frac{\partial Y}{\partial V} \end{pmatrix} = \begin{pmatrix} V & U \\ -V & 1-U \end{pmatrix}$ $|J| = V - VU + VU$
$= V$

$$f_{U,V}(u,v) = \frac{\theta^{\alpha}}{\Gamma(\alpha)} (uv)^{\alpha-1} e^{-\theta uv} \frac{\theta^{\beta}}{\Gamma(\beta)} (v-uv)^{\beta-1} e^{-\theta(v-uv)} \cdot v$$

$$= \frac{\theta^{\alpha+\beta}}{\Gamma(\alpha)\Gamma(\beta)} e^{\cancel{-\theta uv}} \cdot e^{-\theta v} \cdot e^{\cancel{\theta uv}} \cdot u^{\alpha-1} \cdot v^{\alpha-1} \cdot v^{\beta-1} \cdot (1-u)^{\beta-1} \cdot v$$

$$= \frac{\theta^{\alpha+\beta}}{\cancel{\Gamma(\alpha)\Gamma(\beta)}} \cdot \frac{\cancel{\Gamma(\alpha)\Gamma(\beta)}}{\Gamma(\alpha+\beta)} \cdot \underbrace{\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}}_{B^{-1}} \cdot e^{-\theta v} v^{\alpha+\beta-1} \cdot u^{\alpha-1}(1-u)^{\beta-1}$$

$$= \frac{\theta^{\alpha+\beta}}{\Gamma(\alpha+\beta)} v^{\alpha+\beta-1} e^{-\theta v} \cdot \frac{1}{B(\alpha,\beta)} u^{\alpha-1}(1-u)^{\beta-1}$$

So, $U = \frac{X}{X+Y} \sim Beta(\alpha, \beta) = \frac{1}{B(\alpha,\beta)} u^{\alpha-1}(1-u)^{\beta-1}$

3. and $V = X+Y \perp\!\!\!\perp U = \frac{X}{X+Y}$ since the joint distribution of $f_{U,V}(u,v) = f_U(u) \cdot f_V(v)$

4. A2.7. prove Theorem A.2.7

If $Y \sim N_p(u, \Sigma)$, $Var(Y^TAY) = 2trace(A\Sigma A\Sigma) + 4u^TA\Sigma Au$

MGF of Y^TAY

$$M_{Y^TAY}(t) = E(e^{tY^TAY}) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{ty^TAy} K_1 e^{-\frac{(y-u)^T\Sigma^{-1}(y-u)}{2}} dy$$

$$= K_1 \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-[y^T(I-2tA\Sigma)\Sigma^{-1}y - 2u^T\Sigma^{-1}y + u^T\Sigma^{-1}u]/2} dy$$

where $K_1 = \frac{1}{(\sqrt{2\pi})^p |\Sigma|^{\frac{1}{2}}}$, $dy = dy_1 dy_2 \cdots dy_p$.

let $\theta^T = u^T(I-2tA\Sigma)^{-1}$ $V^{-1} = (I-2tA\Sigma)\Sigma^{-1}$

$K_2 = (\sqrt{2\pi})^p |V|^{\frac{1}{2}} e^{-[u^T\Sigma^{-1}u - \theta^TV^{-1}\theta]/2}$

Then $M_{Y^TAY}(t) = k_1 k_2 \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{1}{(\sqrt{2\pi})^p |V|^{\frac{1}{2}}} \cdot e^{-[u^T\Sigma^{-1}u - \theta^T V^{-1}\theta]/2} \, dy}_{\text{the densiy of MVN } = 1}$

$$= |I - 2tA\Sigma|^{-\frac{1}{2}} e^{-u^T(I-(I-2tA\Sigma)^{-1})\Sigma^{-1}u/2}$$

let $C = I - 2tA\Sigma$

$k(t) = \ln[M_{Y^TAY}(t)] = -\frac{1}{2}\ln|C| - \frac{1}{2}u'(I - C^{-1})\Sigma^{-1}u$

$k''(t) = \frac{1}{2}\frac{1}{|C|^2}\left[\frac{d|C|}{dt}\right]^2 - \frac{1}{2}\frac{1}{|C|}\frac{d^2|C|}{dt^2} - \frac{1}{2}u^TC^{-1}\frac{d^2C}{dt^2}C^{-1}\Sigma^{-1}u + u\left[C^{-1}\frac{dC}{dt}\right]^2 C^{-1}\Sigma^{-1}u$

eigenvalue of $A\Sigma$ are λ_i, $i = 1, \cdots, p$

$|C| = \prod_{i=1}^{p}(1 - 2t\lambda_i) = 1 - 2t\sum_i \lambda_i + 4t^2 \sum_{i\neq j}\lambda_i\lambda_j \cdots$

$t = 0 \Rightarrow |C| = 1$

$\frac{d|C|}{dt} = -2\sum_i \lambda_i + 8t\sum_{i\neq j}\lambda_i\lambda_j + O(t)$

$\frac{d^2|C|}{dt^2} = 8\sum_{i\neq j}\lambda_i\lambda_j + O(t)$

$O(t)$ means higher-order terms in t.

when $t \to 0$, $\frac{d|C|}{dt} = -2\,tr(A\Sigma)$ $\quad \frac{d^2|C|}{dt^2} = 8\sum_{i\neq j}\lambda_i\lambda_j$

for $t = 0$, $C = I$, $C^{-1} = I$, $\frac{dC}{dt} = 2A\Sigma$ $\quad \frac{d^2C}{dt} = 0$

So $K''(0) = 2[tr(A\Sigma)]^2 - 4\sum_{i\neq j}\lambda_i\lambda_j + 0 + 4u^TA\Sigma Au$

$$= 2\left\{[tr(A\Sigma)]^2 - 2\sum_{i\neq j}\lambda_i\lambda_j\right\} + 4u^TA\Sigma Au$$

Since $[tr(A)]^2 = tr(A^2) + 2\sum\sum_{i\neq j}\lambda_i\lambda_j$

$$= 2tr((A\Sigma)^2) + 4u^TA\Sigma Au$$

$$= 2tr(A\Sigma A\Sigma) + 4u^TA\Sigma Au$$

5.

A3.3

assume $y_1, \cdots, y_n \overset{iid}{\sim} exp(u)$ $\quad f(y_i) = ue^{-uy_i}$, $y_i \geq 0$

log likelihood contributed by unit i

$$\log(u) - uy_i$$

$\hat{u}$ maximize loglikelihood $\frac{1}{n}\sum_{i=1}^{n}(\log(u) - uy_i)$

$$\frac{\partial \log f(y_i|u)}{\partial u} = \frac{1}{u} - y_i$$

$$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{u} - y_i\right) = 0$$

$$\frac{1}{n}\left(\frac{n}{u} - \sum_{i=1}^{n}y_i\right) = 0$$

$$\frac{\partial^2 \log f(y_i|u)}{\partial u^2} = -\frac{1}{u^2}$$

mle $\hat{u} = \frac{n}{\sum_{i=1}^{n}y_i}$

$$I(\theta) = E\left(-\frac{\partial^2 \log f(y|\theta)}{\partial\theta\partial\theta^T}\right)$$

$$\hat{I}(u) = \frac{1}{n}\sum_{i=1}^{n} -\frac{\partial^2 \log f(y|\theta)}{\partial\theta\partial\theta^T} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{u^2}$$

$$\hat{J}(u) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\partial \log f(y_i|\theta)}{\partial\theta}\right)^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{u}-y_i\right)^2$$

according to lecture

covariance matrix estimator for MLE under misspecification

is $I_n(\hat{\theta})^{-1} J_n(\hat{\theta}) I_n(\hat{\theta})^{-1} = \left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{u^2}\right)^{-1}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{u}-y_i\right)^2\left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{u^2}\right)^{-1}$

If correctly specified

covariance is $\hat{I}(u)^{-1}$