

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

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

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

$$AB_{12} + BB_{22} = 0 \Rightarrow B_{12} = -A^{-1}B B_{22} \quad \text{plug in } AB_{11} + BB_{21} = I \Rightarrow$$
$$CB_{11} + DB_{21} = 0 \Rightarrow B_{21} = -D^{-1}C B_{11} \quad CB_{12} + DB_{22} = I$$

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

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

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

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

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

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

$$\begin{aligned}
(A - BD^{-1}C)^{-1} &= (A(I - A^{-1}BD^{-1}C))^{-1} \\
&= (I - A^{-1}BD^{-1}C)^{-1} A^{-1} \\
&= (I - A^{-1}BD^{-1}C)^{-1} \underbrace{(I - A^{-1}BD^{-1}C + A^{-1}BD^{-1}C) A^{-1}}_{\text{Cancel}} \\
&= (I + (I - A^{-1}BD^{-1}C)^{-1} A^{-1}BD^{-1}C) A^{-1} \\
&= A^{-1} + (I - A^{-1}BD^{-1}C)^{-1} A^{-1}BD^{-1}CA^{-1}
\end{aligned}$$

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  $Y$

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

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

Sherman - Morrison

$$A^{-1} - (I + V^T A^{-1} u)^{-1} A^{-1} u v^T A^{-1} = A^{-1} - \frac{A^{-1} u v^T A^{-1}}{I + V^T A^{-1} u} \quad \text{scalar denominator}$$

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

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

$$= I + u v^T A^{-1} - \frac{u v^T A^{-1} u v^T A^{-1} + A A^{-1} u v^T A^{-1}}{I + V^T A^{-1} u}$$

$$= I + u v^T A^{-1} - \frac{u (I + V^T A^{-1} u) v^T A^{-1}}{I + V^T A^{-1} u}$$

$$= I + u v^T A^{-1} - u v^T A^{-1}$$

$$= I$$

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

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

2. A1.2

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

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  $y = Ax$

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

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

$$= Ax + A^T x$$

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

### 3. A2.1 prove Theorem A2.1

1. let  $Z = X + Y$ ,  $X \sim \text{Gamma}(\alpha, \theta)$ ,  $Y \sim \text{Gamma}(\beta, \theta)$

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

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

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

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

$$x=z \Rightarrow w=1$$

$$= \frac{\theta^{\alpha+\beta} e^{-\theta z}}{\Gamma(\alpha) \Gamma(\beta)} \int_0^1 w^{\alpha-1} z^{\beta-1} (1-\frac{w}{z})^{\beta-1} dz$$

$$= \frac{\theta^{\alpha+\beta} e^{-\theta z}}{\Gamma(\alpha) \Gamma(\beta)} \int_0^1 w^{(\alpha-1)} z^{(\alpha-1)} z^{\beta-1} (1-w)^{\beta-1} dz$$

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

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

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

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

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

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

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

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

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

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

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

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

4. A2.7. prove Theorem A.2.7

If  $Y \sim N(\mu, \Sigma)$ ,  $\text{Var}(Y^T A Y) = 2\text{trace}(A \Sigma A^T) + 4U^T A \Sigma A u$

MGF of  $Y^T A Y$

$$\begin{aligned} M_{Y^T A Y}(t) &= E(e^{t Y^T A Y}) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{t y^T A y} K_1 e^{-\frac{(y-\mu)^T \Sigma^{-1} (y-\mu)}{2}} dy \\ &= K_1 \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\frac{1}{2} y^T (I - 2tA\Sigma)^{-1} y + 2\mu^T \Sigma^{-1} y + \mu^T \Sigma^{-1} \mu} dy \end{aligned}$$

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

$$\text{let } \Theta^T = U^T (I - 2tA\Sigma)^{-1} \quad V^{-1} = (I - 2tA\Sigma)^{-1}$$

$$K_2 = \frac{1}{(2\pi)^p |V|^{\frac{1}{2}}} e^{-\frac{1}{2} \mu^T \Sigma^{-1} \mu - \theta^T V^{-1} \theta / 2}$$

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

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

let  $C = I - 2tA\Sigma$

$$k(t) = \ln [M_{YAY}(t)] = -\frac{1}{2} \ln |C| - \frac{1}{2} u^T (I - C^{-1}) \Sigma^{-1} u$$

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

Eigenvalue of  $A\Sigma$  are  $\lambda_i$ ,  $i=1, \dots, p$

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

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

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

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

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

when  $t > 0$ ,  $\frac{d|C|}{dt} = -2t(A\Sigma) \quad \frac{d^2|C|}{dt^2} = 8 \sum_i \lambda_i \lambda_j$

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

$$S_0 \quad K'(0) = 2 \operatorname{E}[\operatorname{tr}(A\Sigma)]^2 - 4 \sum_{i \neq j} \lambda_i \lambda_j + 0 + 4u^T A \Sigma A u$$

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

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

$$= 2 \operatorname{tr}(A\Sigma)^2 + 4u^T A \Sigma A u$$

$$= 2 \operatorname{tr}(A\Sigma A) + 4u^T A \Sigma A u$$

5.  
A3.)

$$\text{assume } y_1, \dots, y_n \stackrel{\text{iid}}{\sim} \exp(u) \quad f(y_i) = u e^{-uy_i}, \quad y_i \geq 0$$

log likelihood contributed by unit  $i$

$$\log(u) - uy_i$$

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

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

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

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

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

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

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

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

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

according to lecture under correctly specified model.

If we use  $\hat{I}$ ,  $\hat{u} \xrightarrow{d} N(u, \frac{I^{-1}}{n})$

$\frac{1}{n} \hat{I}^{-1} = \frac{1}{n} \frac{n}{\sum_{i=1}^n \hat{u}_i^2} = \frac{1}{\sum_{i=1}^n \hat{u}_i^2}$  as the variance estimator of  $\hat{u}$

under misspecified model, covariance estimator for the mle  $\hat{\alpha}$  is

$$\sqrt{n}(\hat{\alpha} - u) \xrightarrow{d} N(0, \hat{I}(u) J(u) \hat{I}^{-1}(u))$$

$$\hat{u} \xrightarrow{d} N(u, \frac{1}{n} I(u) J(u) I^{-1}(u))$$

$$\frac{1}{n} \hat{I}^{-1} \hat{J} \hat{I}^{-1} = \frac{1}{n} \cdot \frac{1}{\sum_{i=1}^n \hat{u}_i^2} \cdot \frac{1}{n} \sum_{i=1}^n \left( \frac{1}{\hat{u}_i} - y_i \right)^2 \cdot \frac{1}{\sum_{i=1}^n \hat{u}_i^2}$$

$$L(\theta) \propto \theta^n e^{-\lambda \sum y_i}$$

$$L(\theta) \propto n \ln(\theta) - \lambda \sum y_i$$

$$U(\theta) = \frac{n}{\theta} - \lambda \sum y_i$$

$$\theta = \frac{n}{\sum y_i}$$

$$U'(\theta) = -\frac{n}{\theta^2}$$