Matrix Formulas That Is Useful In Least Square Analysis

(You could always swap indexes properly to get similar results)

1 Difference of Quadratic Forms - The Regression Improvement by Introducing a Variate

For invertible $(n_1+n_2+n_3)\times(n_1+n_2+n_3)$ matrix R (no need symmetric), define Q

$$Q = \begin{bmatrix} Q^{(xx)} & Q^{(xy)} & Q^{(xz)} \\ Q^{(yx)} & Q^{(yy)} & Q^{(yz)} \\ Q^{(zx)} & Q^{(zy)} & Q^{(zz)} \end{bmatrix} = R^{-1} = \begin{bmatrix} R^{(xx)} & R^{(xy)} & R^{(xz)} \\ R^{(yx)} & R^{(yy)} & R^{(yz)} \\ R^{(zx)} & R^{(zy)} & R^{(zz)} \end{bmatrix}^{-1}. \tag{1}$$

For any $\begin{bmatrix} \vec{r}^{(x|x)} & \vec{r}^{(x|y)} & \vec{r}^{(x|z)} \end{bmatrix}$ ($n_4\times(n_1+n_2+n_3)$ matrix),
define quadratic forms $\mathrm{SS}_{R[x,z]}$ and $\mathrm{SS}_{R[x,y,z]}$, solutions $\begin{bmatrix} \vec{a}^{(11)} & \vec{a}^{(12)} & \vec{a}^{(13)} \end{bmatrix}$ and $\begin{bmatrix} \vec{a}_T^{(11)} & \vec{a}_T^{(12)} & \vec{a}_T^{(13)} \end{bmatrix}$ as (replace transpose $\cdot^T$ to conjugate transpose $\cdot^H$ for matrix with complex entries)

$$\mathrm{SS}_{R[x,z]} = \begin{bmatrix} \vec{r}^{(x|x)} & \vec{r}^{(x|z)} \end{bmatrix} \begin{bmatrix} R^{(xx)} & R^{(xz)} \\ R^{(zx)} & R^{(zz)} \end{bmatrix}^{-1} \begin{bmatrix} \vec{r}^{(x|x)} & \vec{r}^{(x|z)} \end{bmatrix}^T,$$

$$\mathrm{SS}_{R[x,y,z]} = \begin{bmatrix} \vec{r}^{(x|x)} & \vec{r}^{(x|y)} & \vec{r}^{(x|z)} \end{bmatrix} \begin{bmatrix} R^{(xx)} & R^{(xy)} & R^{(xz)} \\ R^{(yx)} & R^{(yy)} & R^{(yz)} \\ R^{(zx)} & R^{(zy)} & R^{(zz)} \end{bmatrix}^{-1} \begin{bmatrix} \vec{r}^{(x|x)} & \vec{r}^{(x|y)} & \vec{r}^{(x|z)} \end{bmatrix}^T,$$

$$\begin{bmatrix} \vec{a}^{(11)} & \vec{a}^{(12)} & \vec{a}^{(13)} \end{bmatrix} \begin{bmatrix} R^{(xx)} & R^{(xy)} & R^{(xz)} \\ R^{(yx)} & R^{(yy)} & R^{(yz)} \\ R^{(zx)} & R^{(zy)} & R^{(zz)} \end{bmatrix} = \begin{bmatrix} \vec{r}^{(x|x)} & \vec{r}^{(x|y)} & \vec{r}^{(x|z)} \end{bmatrix},$$

$$\begin{bmatrix} \vec{a}_T^{(11)} & \vec{a}_T^{(12)} & \vec{a}_T^{(13)} \end{bmatrix} \begin{bmatrix} R^{(xx)} & R^{(xy)} & R^{(xz)} \\ R^{(yx)} & R^{(yy)} & R^{(yz)} \\ R^{(zx)} & R^{(zy)} & R^{(zz)} \end{bmatrix}^T = \begin{bmatrix} \vec{r}^{(x|x)} & \vec{r}^{(x|y)} & \vec{r}^{(x|z)} \end{bmatrix}.$$

Then the difference of the quadratic forms is
(Assume the inverse matries in the expressions are all valid. Pseudo-inverse might also work?)

$$\mathrm{SS}_{R[x,y,z]} - \mathrm{SS}_{R[x,z]} = \vec{a}^{(12)} \left(Q^{(yy)}\right)^{-1} \left(\vec{a}_T^{(12)}\right)^T. \tag{2}$$

The $\left(Q^{(yy)}\right)^{-1}$ might be expressed in R blocks as
(Swap index x and z will give another expression)

$$\begin{aligned} \left(Q^{(yy)}\right)^{-1} &= R^{(yy)} - R^{(yx)} \left(R^{(xx)}\right)^{-1} R^{(xy)} \\ &\quad - \left(R^{(yz)} - R^{(yx)} \left(R^{(xx)}\right)^{-1} R^{(xz)}\right) \left(R^{(zz)} - R^{(zx)} \left(R^{(xx)}\right)^{-1} R^{(xz)}\right)^{-1} \left(R^{(zy)} - R^{(zx)} \left(R^{(xx)}\right)^{-1} R^{(xy)}\right) \end{aligned} \tag{3}$$

2 Difference of Solutions - The Correction to Coefficients When Introducing a Variate

If we solve

$$\begin{bmatrix} \vec{b}^{(11)} & \vec{b}^{(12)} \end{bmatrix} \begin{bmatrix} R^{(xx)} & R^{(xy)} \\ R^{(yx)} & R^{(yy)} \end{bmatrix} = \begin{bmatrix} \vec{r}^{(x|x)} & \vec{r}^{(x|y)} \end{bmatrix}$$

Then there is relation
(There are 6 of these relations)

$$b^{(12)} = a^{(12)} - a^{(13)} \left(Q^{(zz)}\right)^{-1} Q^{(zy)} \tag{4}$$

$$a^{(12)} \left(I - \left(Q^{(yy)}\right)^{-1} Q^{(yz)} \left(Q^{(zz)}\right)^{-1} Q^{(zy)}\right) = b^{(12)} + b^{(13)}_{[xz]} \left(Q^{(zz)}\right)^{-1} Q^{(zy)} \tag{5}$$

Define P

$$P = \begin{bmatrix} P^{(xx)} & P^{(xy)} \\ P^{(yx)} & P^{(yy)} \end{bmatrix} \triangleq \begin{bmatrix} R^{(xx)} & R^{(xy)} \\ R^{(yx)} & R^{(yy)} \end{bmatrix}^{-1}.$$

There are relations between P, Q and R

$$P = \begin{bmatrix} Q^{(xx)} & Q^{(xy)} \\ Q^{(yx)} & Q^{(yy)} \end{bmatrix} - \begin{bmatrix} Q^{(xz)} \\ Q^{(yz)} \end{bmatrix} \left(Q^{(zz)}\right)^{-1} \begin{bmatrix} Q^{(zx)} & Q^{(zy)} \end{bmatrix} \tag{6}$$

$$\left(Q^{(zz)}\right)^{-1} = R^{(zz)} - \begin{bmatrix} R^{(zx)} & R^{(zy)} \end{bmatrix} \begin{bmatrix} R^{(xx)} & R^{(xy)} \\ R^{(yx)} & R^{(yy)} \end{bmatrix}^{-1} \begin{bmatrix} R^{(xz)} \\ R^{(yz)} \end{bmatrix}$$

$$\begin{aligned} \left(Q^{(zz)}\right)^{-1} Q^{(zy)} &= -\left(R^{(zy)} - R^{(zx)} \left(R^{(xx)}\right)^{-1} R^{(xy)}\right) \left(R^{(yy)} - R^{(yx)} \left(R^{(xx)}\right)^{-1} R^{(xy)}\right)^{-1} \\ &= -\begin{bmatrix} R^{(zx)} & R^{(zy)} \end{bmatrix} \begin{bmatrix} P^{(xy)} \\ P^{(yy)} \end{bmatrix} \end{aligned}$$

$$I - \left(Q^{(yy)}\right)^{-1} Q^{(yz)} \left(Q^{(zz)}\right)^{-1} Q^{(zy)} = \left(Q^{(yy)}\right)^{-1} P^{(yy)}$$

Sometimes $\vec{b}^{(11)}$ is also wanted

$$\vec{b}^{(11)} = \vec{a}^{(11)} - a^{(13)} \left(Q^{(zz)}\right)^{-1} Q^{(zx)}$$

In matrix form

$$\begin{bmatrix} \vec{b}^{(11)} & \vec{b}^{(12)} \\ \vec{b}^{(21)} & \vec{b}^{(22)} \end{bmatrix} = \begin{bmatrix} \vec{a}^{(11)} & \vec{a}^{(12)} \\ \vec{a}^{(21)} & \vec{a}^{(22)} \end{bmatrix} - \begin{bmatrix} a^{(13)} \\ a^{(23)} \end{bmatrix} \left(Q^{(zz)}\right)^{-1} \begin{bmatrix} Q^{(zx)} & Q^{(zy)} \end{bmatrix}$$

$$\begin{bmatrix} \vec{b}^{(11)} & \vec{b}^{(12)} \\ \vec{b}^{(21)} & \vec{b}^{(22)} \end{bmatrix} = \begin{bmatrix} \vec{a}^{(11)} & \vec{a}^{(12)} \\ \vec{a}^{(21)} & \vec{a}^{(22)} \end{bmatrix} + \begin{bmatrix} a^{(13)} \\ a^{(23)} \end{bmatrix} \begin{bmatrix} R^{(zx)} & R^{(zy)} \end{bmatrix} \begin{bmatrix} R^{(xx)} & R^{(xy)} \\ R^{(yx)} & R^{(yy)} \end{bmatrix}^{-1}$$

$$\begin{bmatrix} a^{(13)} \\ a^{(23)} \end{bmatrix} \begin{bmatrix} R^{(zx)} & R^{(zy)} \end{bmatrix} = \begin{bmatrix} \vec{r}^{(x|x)} & \vec{r}^{(x|y)} \\ \vec{r}^{(y|x)} & \vec{r}^{(y|y)} \end{bmatrix} - \begin{bmatrix} \vec{a}^{(11)} & \vec{a}^{(12)} \\ \vec{a}^{(21)} & \vec{a}^{(22)} \end{bmatrix} \begin{bmatrix} R^{(xx)} & R^{(xy)} \\ R^{(yx)} & R^{(yy)} \end{bmatrix}$$

$$\begin{bmatrix} a^{(13)} \\ a^{(23)} \end{bmatrix} \begin{bmatrix} R^{(zx)} & R^{(zy)} \end{bmatrix} \begin{bmatrix} R^{(xx)} & R^{(xy)} \\ R^{(yx)} & R^{(yy)} \end{bmatrix}^{-1} = \begin{bmatrix} \vec{r}^{(x|x)} & \vec{r}^{(x|y)} \\ \vec{r}^{(y|x)} & \vec{r}^{(y|y)} \end{bmatrix} \begin{bmatrix} R^{(xx)} & R^{(xy)} \\ R^{(yx)} & R^{(yy)} \end{bmatrix}^{-1} - \begin{bmatrix} \vec{a}^{(11)} & \vec{a}^{(12)} \\ \vec{a}^{(21)} & \vec{a}^{(22)} \end{bmatrix}$$

3 Relation to Linear Regression and Granger Causality For Time Series

Assume $\mathrm{E}(x_t) = \mathrm{E}(y_t) = \mathrm{E}(y_t) = O$, as usual, let's do a 3-var AR of order m:

$$\begin{cases} x_t = \sum_{j=1}^{m} a_j^{(11)} x_{t-j} + \sum_{j=1}^{m} a_j^{(12)} y_{t-j} + \sum_{j=1}^{m} a_j^{(13)} z_{t-j} + \epsilon_t^{(1)} & \text{(a)} \\ y_t = \sum_{j=1}^{m} a_j^{(21)} x_{t-j} + \sum_{j=1}^{m} a_j^{(22)} y_{t-j} + \sum_{j=1}^{m} a_j^{(23)} z_{t-j} + \epsilon_t^{(2)} & \text{(b)} \\ z_t = \sum_{j=1}^{m} a_j^{(31)} x_{t-j} + \sum_{j=1}^{m} a_j^{(32)} y_{t-j} + \sum_{j=1}^{m} a_j^{(33)} z_{t-j} + \epsilon_t^{(3)} & \text{(c)} \end{cases} \tag{7}$$

The least square solution for $\vec{a}^{(11)} \quad \vec{a}^{(12)} \quad \vec{a}^{(13)}$ in min var $\epsilon_t^{(1)}$ $(\mathrm{E}(\epsilon_t^{(u)} x_{t-k}^{(v)T}) = 0,\ \forall k = 1 \ldots m)$ satisfies

$$\begin{bmatrix} \vec{a}^{(11)} & \vec{a}^{(12)} & \vec{a}^{(13)} \end{bmatrix} \begin{bmatrix} R^{(xx)} & R^{(xy)} & R^{(xz)} \\ R^{(yx)} & R^{(yy)} & R^{(yz)} \\ R^{(zx)} & R^{(zy)} & R^{(zz)} \end{bmatrix} = \begin{bmatrix} \vec{r}^{(x|x)} & \vec{r}^{(x|y)} & \vec{r}^{(x|z)} \end{bmatrix}, \tag{8}$$

where (here $\mathbb{E}$ can be expectation operator or $\frac{1}{L}\sum_{t=1}^{L}$)

$$\vec{a}^{(uv)} = \left(a_j^{(uv)}\right)_{j=1\ldots m}^{\text{row}}, \quad (u, v \in \{1,2,3\}),$$

$$R^{(uv)} = \left(\mathbb{E}(x_{t-j}^{(u)} x_{t-k}^{(v)T})\right)_{j,k=1\ldots m}, \quad (u, v \in \{x,y,z\}),$$

$$\vec{r}^{(u|v)} = \left(\mathbb{E}(x_t^{(u)} x_{t-k}^{(v)T})\right)_{k=1\ldots m}^{\text{row}}, \quad (u, v \in \{x,y,z\}).$$

$(a_j)_{j=1\ldots m}^{\text{row}}$ means tiling a_j in a row: $(a_j)_{j=1\ldots m}^{\text{row}} = \begin{bmatrix} a_1 & \cdots & a_m \end{bmatrix}$, and $(a_{jk})_{j,k=1\ldots m}$ is a matrix which its j row k column is a_{jk}.

The residual variance of Eq. (7) can be obtained by

$$\mathrm{var}(x_t) = \vec{a}^{(11)} \vec{r}^{(x|x)T} + \vec{a}^{(12)} \vec{r}^{(x|y)T} + \vec{a}^{(13)} \vec{r}^{(x|z)T} + \mathrm{var}(\epsilon_t^{(1)})$$

or

$$\mathrm{var}(\epsilon_t^{(1)}) = \mathrm{var}(x_t) - \begin{bmatrix} \vec{r}^{(x|x)} & \vec{r}^{(x|y)} & \vec{r}^{(x|z)} \end{bmatrix} \begin{bmatrix} R^{(xx)} & R^{(xy)} & R^{(xz)} \\ R^{(yx)} & R^{(yy)} & R^{(yz)} \\ R^{(zx)} & R^{(zy)} & R^{(zz)} \end{bmatrix}^{-1} \begin{bmatrix} \vec{r}^{(x|x)} & \vec{r}^{(x|y)} & \vec{r}^{(x|z)} \end{bmatrix}^T. \tag{9}$$

Similarly for least square min var $\epsilon_t^{(1|3)}$ in (regression of x without y)

$$x_t = \sum_{j=1}^{m} c_j^{(11)} x_{t-j} + \sum_{j=1}^{m} c_j^{(13)} z_{t-j} + \epsilon_t^{(1|3)}, \tag{10}$$

that is

$$\mathrm{var}(\epsilon_t^{(1|3)}) = \mathrm{var}(x_t) - \begin{bmatrix} \vec{r}^{(x|x)} & \vec{r}^{(x|z)} \end{bmatrix} \begin{bmatrix} R^{(xx)} & R^{(xz)} \\ R^{(zx)} & R^{(zz)} \end{bmatrix}^{-1} \begin{bmatrix} \vec{r}^{(x|x)} & \vec{r}^{(x|z)} \end{bmatrix}^T. \tag{11}$$

Conditional Granger causality $F_{y\to x|z}$ is defined as

$$F_{y\to x|z} = \ln \frac{\mathrm{var}(\epsilon_t^{(1|3)})}{\mathrm{var}(\epsilon_t^{(1)})} = \ln\left(1 + \frac{\mathrm{var}(\epsilon_t^{(1|3)}) - \mathrm{var}(\epsilon_t^{(1)})}{\mathrm{var}(\epsilon_t^{(1)})}\right). \tag{12}$$

With help of Eq.(2), we get

$$F_{y\to x|z} = \ln\left(1 + \frac{\vec{a}^{(12)} \left(Q^{(yy)}\right)^{-1} \left(\vec{a}^{(12)}\right)^T}{\mathrm{var}(\epsilon_t^{(1)})}\right) \tag{13}$$

3.1 Conditional GC and pairwise GC

Due to expression Eq.(13), the pairwise GC can be expressed as

$$F_{y\to x|z} = \ln\left(1 + \frac{\vec{b}^{(12)}\left(P^{(yy)}\right)^{-1}\left(\vec{b}^{(12)}\right)^T}{\mathrm{var}(\epsilon_t^{(1|2)})}\right) \tag{14}$$

where (recall Eq.(4) (2) and (6))

$$b^{(12)} = a^{(12)} - a^{(13)}\left(Q^{(zz)}\right)^{-1}Q^{(zy)},$$

$$\mathrm{var}(\epsilon_t^{(1|2)}) = \mathrm{var}(\epsilon_t^{(1)}) + \vec{a}^{(13)}\left(Q^{(zz)}\right)^{-1}\left(\vec{a}^{(13)}\right)^T,$$

$$P^{(yy)} = Q^{(yy)} - Q^{(yz)}\left(Q^{(zz)}\right)^{-1}Q^{(zy)}.$$

The related regression is

$$\begin{cases} x_t = \sum_{j=1}^{m} b_j^{(11)} x_{t-j} + \sum_{j=1}^{m} b_j^{(12)} y_{t-j} + \epsilon_t^{(1|2)} & \text{(a)} \\ y_t = \sum_{j=1}^{m} b_j^{(21)} x_{t-j} + \sum_{j=1}^{m} b_j^{(22)} y_{t-j} + \epsilon_t^{(2|1)} & \text{(b)} \end{cases} \tag{15}$$

From Eq.(14), one can know that the pairwise GC can be expressed by joint regression coefficients and residuals, and Q matrix.

3.2 Approximation of GC

Due to Taylor expansion, we have approximation

$$\ln(1+x) = \sum_{k=1}^{\infty}\frac{(-1)^{k-1}x^k}{k} \approx x, \quad \text{for } x \ll 1. \tag{16}$$

From Eq. (12) get

$$F_{y\to x|z} \approx \frac{\vec{a}^{(12)}\left(Q^{(yy)}\right)^{-1}\left(\vec{a}^{(12)}\right)^T}{\mathrm{var}(\epsilon_t^{(1)})}. \tag{17}$$

For whitened signal and a weak GC, we might even use to get a reasonable approximation

$$F_{y\to x|z} \approx \frac{\mathrm{var}\,(\epsilon_t^{(2)})}{\mathrm{var}\,(\epsilon_t^{(1)})}\vec{a}^{(12)}\vec{a}^{(12)T} \tag{18}$$

3.3 Partial correlation coefficient (seems useless)

$$\begin{bmatrix} \vec{a}^{(11)} & \vec{a}^{(12)} & \vec{a}^{(13)} \end{bmatrix}\begin{bmatrix} R^{(xx)} & R^{(xy)} & O \\ R^{(yx)} & R^{(yy)} & O \\ O & O & R^{(zz)} \end{bmatrix} = \begin{bmatrix} \vec{r}^{(x|x)} & \vec{r}^{(x|y)} & 0 \end{bmatrix}$$

$$\left(Q^{(yy)}\right)^{-1} = R^{(yy)} - R^{(yx)}\left(R^{(xx)}\right)^{-1}R^{(xy)}$$

$$x_t' = x_t - \alpha Y$$

What if z is totally whitened? i.e.

$$R^{(zz)} = I$$

not verified

$$c^{(33)} = a^{(33)} - \begin{bmatrix} a^{(31)} & a^{(32)} \end{bmatrix}\begin{bmatrix} Q^{(xx)} & Q^{(xy)} \\ Q^{(yx)} & Q^{(yy)} \end{bmatrix}^{-1}\begin{bmatrix} Q^{(xz)} \\ Q^{(yz)} \end{bmatrix}$$

4 Spectrum Domain Approximation

From Eq. (4) get analog formula for frequency domain:

$$B_{xy} \approx A_{xy} - \left[A_{xz} Q_{zz}^{-1} Q_{zy}\right]_+$$

where $[\cdot]_+$ stand for extract the positive time part. That is

$$[f(w)]_+ = \mathscr{F}_{\mathrm{DFT}}\left[\mathscr{F}_{\mathrm{DFT}}^{-1}\left[f(w), t\right](t) \cdot u(t), w\right]$$

$$u(t) = \begin{cases} 1 & t \geq 0 \\ 0 & t < 0 \end{cases}$$

Under assumption (blocked diagonal matrix)

$$\Sigma = \mathrm{diag}\left(\left[\begin{array}{ccc} \Sigma_{xx} & \Sigma_{yy} & \Sigma_{zz} \end{array}\right]\right)$$

We have

$$Q_{zz} = A_{xz}^{*T} \Sigma_{xx}^{-1} A_{xz} + A_{yz}^{*T} \Sigma_{yy}^{-1} A_{yz} + A_{zz}^{*T} \Sigma_{zz}^{-1} A_{zz}$$
$$Q_{zy} = A_{xz}^{*T} \Sigma_{xx}^{-1} A_{xy} + A_{yz}^{*T} \Sigma_{yy}^{-1} A_{yy} + A_{zz}^{*T} \Sigma_{zz}^{-1} A_{zy}$$

Hence

$$A_{xz}/Q_{zz} * Q_{zy} \approx A_{xz}\left(A_{zz}^{-1} \Sigma_{zz} \left(A_{zz}^{*T}\right)^{-1}\right)\left(A_{yz}^{*T} \Sigma_{yy}^{-1} A_{yy} + A_{zz}^{*T} \Sigma_{zz}^{-1} A_{zy}\right) \approx A_{xz}\left(A_{zy} + A_{yz}^*\right)$$

$$B_{xy} \approx A_{xy} - \left[A_{xz}\left(A_{zy} + A_{yz}^*\right)\right]_+$$

$$Q_{zz}^{-1} = S_{zz} - S_{z,xy} S_{xy,xy}^{-1} S_{xy,z}$$

——————-

$$\left[\begin{array}{ccc} O & A_{xy} & A_{xz} \\ & & \end{array}\right]\left(\left[\begin{array}{c} A_{xy} \\ O \\ A_{zy} \end{array}\right] + \left[\begin{array}{ccc} A_{yz} & O & A_{yz} \end{array}\right]^{*T}\right) = A(A + A^*)$$

as

5 Approximation of Toeplitz Matrix Operations

Mul of band toeplitz matrix

$$\left[\begin{array}{ccc} a_0 & a_1 & \\ a_{-1} & a_0 & a_1 \\ & a_{-1} & a_0 \end{array}\right]\left[\begin{array}{ccc} b_0 & b_1 & \\ b_{-1} & b_0 & b_1 \\ & b_{-1} & b_0 \end{array}\right] = \left[\begin{array}{ccc} c_0 - a_{-1}b_1 & c_1 & c_2 \\ c_{-1} & c_0 & c_1 \\ c_{-2} & c_{-1} & c_0 - a_1 b_{-1} \end{array}\right] \tag{19}$$

where $c_0 = a_{-1}b_1 + a_0 b_0 + a_1 b_{-1}$, $c_{-1} = a_{-1}b_0 + a_0 b_{-1}$, $c_1 = a_0 b_1 + a_1 b_0$.

Inverse of band toeplitz ??????????

$$R(Q + G) = I$$

$$GR = I - QR$$

$$GRQ = Q - QRQ$$

since

$$RQ \approx I$$

so

$$G \approx Q - QRQ$$

$$R^{-1} \approx Q + Q - QRQ$$

In summary, that is

$$K = I - QR$$

$$R^{-1} = (I - K)^{-1} Q$$

Bad......since it possible that $\|K\| > 1$. Indeed, there is a 10 neuron case that $\|K\| = 8.34$ (for non-whitenned covz). But seems work for whitened signal and $p < 40$.

6 Appendix: Formulas Used to Derive Above Results

6.1 Binomial inverse theorem

$$(A+UBV)^{-1} = A^{-1} - A^{-1}UB\left(B + BVA^{-1}UB\right)^{-1}BVA^{-1}$$

Specially

$$\left(A - BD^{-1}C\right)^{-1} = A^{-1} + A^{-1}B\left(D - CA^{-1}B\right)^{-1}CA^{-1}$$

6.2 Inverse of $(n_1+n_2)\times(n_1+n_2)$ matrix

(Assume the inverse matries in the expressions are all valid)

$$\left(R^{(zy)} - R^{(zx)}\left(R^{(xx)}\right)^{-1}R^{(xy)}\right)\left(R^{(yy)} - R^{(yx)}\left(R^{(xx)}\right)^{-1}R^{(xy)}\right)^{-1}$$

Use row elimination, get

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \left(a_{11} - a_{12}a_{22}^{-1}a_{21}\right)^{-1} & -\left(a_{11} - a_{12}a_{22}^{-1}a_{21}\right)^{-1}a_{12}a_{22}^{-1} \\ -\left(a_{22} - a_{21}a_{11}^{-1}a_{12}\right)^{-1}a_{21}a_{11}^{-1} & \left(a_{22} - a_{21}a_{11}^{-1}a_{12}\right)^{-1} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}^{-1} + a_{11}^{-1}a_{12}\left(a_{22} - a_{21}a_{11}^{-1}a_{12}\right)^{-1}a_{21}a_{11}^{-1} & -a_{11}^{-1}a_{12}\left(a_{22} - a_{21}a_{11}^{-1}a_{12}\right)^{-1} \\ -a_{22}^{-1}a_{21}\left(a_{11} - a_{12}a_{22}^{-1}a_{21}\right)^{-1} & a_{22}^{-1} + a_{22}^{-1}a_{21}\left(a_{11} - a_{12}a_{22}^{-1}a_{21}\right)^{-1}a_{12}a_{22}^{-1} \end{bmatrix}$$

6.2.1 Known inverse matrix, get inverse of sub matrix

Known

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{-1} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

Then a_{22}^{-1} can get from

$$a_{22}^{-1} = b_{22} - b_{21}\left(b_{11}\right)^{-1}b_{12}$$

6.3 Inverse of $(n_1+n_2+n_3)\times(n_1+n_2+n_3)$ matrix

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}^{-1} = UD^{-1}L$$

$$U = \begin{bmatrix} I & -a_{11}^{-1}a_{12} & a_{11}^{-1}a_{12}\left(a_{22} - a_{21}a_{11}^{-1}a_{12}\right)^{-1}\left(a_{23} - a_{21}a_{11}^{-1}a_{13}\right) - a_{11}^{-1}a_{13} \\ O & I & -\left(a_{22} - a_{21}a_{11}^{-1}a_{12}\right)^{-1}\left(a_{23} - a_{21}a_{11}^{-1}a_{13}\right) \\ O & O & I \end{bmatrix}$$

$$L = \begin{bmatrix} I & O & O \\ -a_{21}a_{11}^{-1} & I & O \\ \left(a_{32} - a_{31}a_{11}^{-1}a_{12}\right)\left(a_{22} - a_{21}a_{11}^{-1}a_{12}\right)^{-1}a_{21}a_{11}^{-1} - a_{31}a_{11}^{-1} & -\left(a_{32} - a_{31}a_{11}^{-1}a_{12}\right)\left(a_{22} - a_{21}a_{11}^{-1}a_{12}\right)^{-1} & I \end{bmatrix}$$

$$D = \begin{bmatrix} a_{11} & O & O \\ O & a_{22} - a_{21}a_{11}^{-1}a_{12} & O \\ O & O & a_{33} - a_{31}a_{11}^{-1}a_{13} - \left(a_{32} - a_{31}a_{11}^{-1}a_{12}\right)\left(a_{22} - a_{21}a_{11}^{-1}a_{12}\right)^{-1}\left(a_{23} - a_{21}a_{11}^{-1}a_{13}\right) \end{bmatrix}$$

Note:

$$\begin{bmatrix} 1 & a_{12} & a_{13} \\ 0 & 1 & a_{23} \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} b_1 & 0 & 0 \\ 0 & b_2 & 0 \\ 0 & 0 & b_3 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ c_{21} & 1 & 0 \\ c_{31} & c_{32} & 1 \end{bmatrix} = \begin{bmatrix} b_1 + a_{12}b_2c_{21} + a_{13}b_3c_{31} & a_{12}b_2 + a_{13}b_3c_{32} & a_{13}b_3 \\ b_2c_{21} + a_{23}b_3c_{31} & b_2 + a_{23}b_3c_{32} & a_{23}b_3 \\ b_3c_{31} & b_3c_{32} & b_3 \end{bmatrix}$$

7 Useful Theorems

7.1 Spectral Representation of Multivariate Stationary Time Series [1]

$$R(h) = \int_{-\pi}^{\pi} e^{ihw} \mathrm{d}F(w), \quad h = 0, \pm 1, \ldots$$

$F(-\pi) = 0$, F is right-continuous, $F(\mu) - F(\lambda)$ non-negative define for all $\lambda < \mu$.

- $E\left(Z(\lambda)Z^*(\lambda)\right)$ finite
- $E\left(Z(\lambda)\right) = 0$, $-\pi < \lambda < \pi$
- $E\left(Z(\lambda_4) - Z(\lambda_3)\right)\left(Z(\lambda_2) - Z(\lambda_1)\right)^* = 0$ $(\lambda_1, \lambda_2] \cap (\lambda_3, \lambda_4] = \phi$
- $E\left(Z(\lambda + \delta) - Z(\lambda)\right)\left(Z(\lambda + \delta) - Z(\lambda)\right)^* \to 0$, as $\delta \downarrow 0$.

unique matrix distribution F, $F(\mu) - F(\lambda) = E\left(\left(Z(\mu) - Z(\lambda)\right)\left(Z(\mu) - Z(\lambda)\right)^*\right)$, $\lambda \leq \mu$. o

$$E\left(\mathrm{d}Z(\lambda)\mathrm{d}Z^*(\lambda)\right) = \begin{cases} \mathrm{d}F(\lambda) & \mu = \lambda \\ 0 & \text{otherwise} \end{cases}$$

$$X_t = \int_{-\pi}^{\pi} e^{itv} \mathrm{d}Z(v), \quad t = 0, \pm 1, \ldots \quad \text{with probability 1}$$

References

[1] Peter J. Brockwell and Richard A. Davis. *Time Series: Theory and Methods*. Springer Science & Business Media, November 2013.