

Matrix Formulas That Is Useful In Least Square Analysis

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

1 Difference of Quadratic Forms, Inverse Matrices and Solution of Linear Systems

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

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

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 $SS_{R[x,z]}$, $SS_{R[x,y,z]}$, $\begin{bmatrix} \vec{a}^{(11)} & \vec{a}^{(12)} & \vec{a}^{(13)} \end{bmatrix}$ 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)

$$\begin{aligned} SS_{R[x,z]} &= \begin{bmatrix} \vec{r}^{(x|x)} & \vec{r}^{(x|z)} \end{bmatrix} \begin{bmatrix} R^{(xx)} & R^{(xz)} \\ R^{(zx)} & R^{(zz)} \end{bmatrix}^{-1} \begin{bmatrix} \vec{r}^{(x|x)} & \vec{r}^{(x|z)} \end{bmatrix}^T, \\ SS_{R[x,y,z]} &= \begin{bmatrix} \vec{r}^{(x|x)} & \vec{r}^{(x|y)} & \vec{r}^{(x|z)} \end{bmatrix} \begin{bmatrix} R^{(xx)} & R^{(xy)} & R^{(xz)} \\ R^{(yx)} & R^{(yy)} & R^{(yz)} \\ R^{(zx)} & R^{(zy)} & R^{(zz)} \end{bmatrix}^{-1} \begin{bmatrix} \vec{r}^{(x|x)} & \vec{r}^{(x|y)} & \vec{r}^{(x|z)} \end{bmatrix}^T, \\ \begin{bmatrix} \vec{a}^{(11)} & \vec{a}^{(12)} & \vec{a}^{(13)} \end{bmatrix} \begin{bmatrix} R^{(xx)} & R^{(xy)} & R^{(xz)} \\ R^{(yx)} & R^{(yy)} & R^{(yz)} \\ R^{(zx)} & R^{(zy)} & R^{(zz)} \end{bmatrix} &= \begin{bmatrix} \vec{r}^{(x|x)} & \vec{r}^{(x|y)} & \vec{r}^{(x|z)} \end{bmatrix}, \\ \begin{bmatrix} \vec{a}_T^{(11)} & \vec{a}_T^{(12)} & \vec{a}_T^{(13)} \end{bmatrix} \begin{bmatrix} R^{(xx)} & R^{(xy)} & R^{(xz)} \\ R^{(yx)} & R^{(yy)} & R^{(yz)} \\ R^{(zx)} & R^{(zy)} & R^{(zz)} \end{bmatrix}^T &= \begin{bmatrix} \vec{r}^{(x|x)} & \vec{r}^{(x|y)} & \vec{r}^{(x|z)} \end{bmatrix}. \end{aligned}$$

Then the difference of the quadratic form is

(Assume the inverse matrices in the expressions are all valid, pseudo-inverse might also work)

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

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

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

Define P

$$\begin{bmatrix} P^{(xx)} & P^{(xy)} \\ P^{(yx)} & P^{(yy)} \end{bmatrix} \triangleq \begin{bmatrix} R^{(xx)} & R^{(xy)} \\ R^{(yx)} & R^{(yy)} \end{bmatrix}^{-1} = \begin{bmatrix} Q^{(xx)} & Q^{(xy)} \\ Q^{(yx)} & Q^{(yy)} \end{bmatrix} - \begin{bmatrix} Q^{(xz)} \\ Q^{(yz)} \end{bmatrix} \left(Q^{(zz)} \right)^{-1} \begin{bmatrix} Q^{(zx)} & Q^{(zy)} \end{bmatrix} \quad (4)$$

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

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)

$$b^{(12)} = a^{(12)} - a^{(13)} \left(Q^{(zz)} \right)^{-1} Q^{(zy)} \quad (5)$$

$$a^{(12)} \left(I - \left(Q^{(yy)} \right)^{-1} Q^{(yz)} \left(Q^{(zz)} \right)^{-1} Q^{(zy)} \right) = b^{(12)} + b_{[xz]}^{(13)} \left(Q^{(zz)} \right)^{-1} Q^{(zy)} \quad (6)$$

Note

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

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

Also

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

In matrix form

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

Q: partial coef and coef??—should correct.

1.1 Partial correlation coefficient

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

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

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

What if z is totally whitened? i.e.

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

not verified

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

1.2 Spectrum domain approximation

From Eq. (5) get analog formula for frequency domain:

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

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

$$[f(w)]_+ = \mathcal{F}_{\text{DFT}}^{-1} [\mathcal{F}_{\text{DFT}} [f(w), t](t) \cdot u(t), w]$$

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

Under assumption (blocked diagonal matrix)

$$\Sigma = \text{diag} \left(\begin{bmatrix} \Sigma_{xx} & \Sigma_{yy} & \Sigma_{zz} \end{bmatrix} \right)$$

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

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

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

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

$$\begin{bmatrix} O & A_{xy} & A_{xz} \end{bmatrix} \left(\begin{bmatrix} A_{xy} \\ O \\ A_{zy} \end{bmatrix} + \begin{bmatrix} A_{yz} & O & A_{yz} \end{bmatrix}^{*T} \right) = A(A + A^*)$$

as

2 Approximation of Toeplitz Matrix Operations

Mul of band toeplitz matrix

$$\begin{bmatrix} a_0 & a_1 & \\ a_{-1} & a_0 & a_1 \\ & a_{-1} & a_0 \end{bmatrix} \begin{bmatrix} b_0 & b_1 & \\ b_{-1} & b_0 & b_1 \\ & b_{-1} & b_0 \end{bmatrix} = \begin{bmatrix} c_0 - a_{-1}b_1 & c_1 & c_2 \\ c_{-1} & c_0 & c_1 \\ c_{-2} & c_{-1} & c_0 - a_1b_{-1} \end{bmatrix} \quad (7)$$

where $c_0 = a_{-1}b_1 + a_0b_0 + a_1b_{-1}$, $c_{-1} = a_{-1}b_0 + a_0b_{-1}$, $c_1 = a_0b_1 + a_1b_0$.

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 $p < 40$.

3 Appendix: Formulas Used to Derive Above Results

3.1 Binomial inverse theorem

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

Specially

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

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

(Assume the inverse matrices 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

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

3.2.1 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

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

3.3 Inverse of $(n_1 + n_2 + n_3) \times (n_1 + n_2 + n_3)$ maxtrix

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

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

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

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

Note:

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

4 Useful Theorems

4.1 Spectral Representation of Multivariate Stationary Time Series [1]

$$R(h) = \int_{-\pi}^{\pi} e^{ihw} dF(w), \quad h = 0, \pm 1, \dots$$

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

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

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

$$E(dZ(\lambda)dZ^*(\lambda)) = \begin{cases} dF(\lambda) & \mu = \lambda \\ 0 & \text{otherwise} \end{cases}$$

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

References

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