

---

# MATRIX APPROCH TO GRS STATISTIC

---

Xie zejian

xiezej@gmail.com

November 11, 2020

Suppose we run time series regressions for each of the  $n$  risky assets

$$\mathbf{r}_t^e = \alpha + \beta r_{m,t}^e + \nu_t$$

where  $\alpha, \mathbf{r}_t^e, \beta, \nu_t$  are  $n \times 1$  vector and  $r_{m,t}^e$  is scalar.

By the discussion above,  $\alpha = \mathbf{0}$  when CAPM holds. Assume  $\{\nu_t\}_{t=1}^T$  i.i.d with  $\mathcal{N}(0, \Sigma)$ , we have  $\mathbf{r}_t^e \mid r_{m,t}^e \sim \mathcal{N}(\alpha + \beta r_{m,t}^e, \Sigma)$ .

Suppose

$$\mathbf{r} = \begin{bmatrix} \mathbf{r}_1^{e'} \\ \mathbf{r}_2^{e'} \\ \mathbf{r}_3^{e'} \\ \vdots \\ \mathbf{r}_T^{e'} \end{bmatrix}, \mathbf{r}_m = \begin{bmatrix} r_{m,1}^e \\ r_{m,2}^e \\ r_{m,3}^e \\ \vdots \\ r_{m,T}^e \end{bmatrix}$$

Now  $\mathbf{r}' \sim \mathcal{MN}_{n \times T}(\alpha \mathbf{e}' + \beta \mathbf{r}_m', \Sigma, \mathbf{I})$ , the p.d.f is (Wikipedia contributors 2019)

$$p(\mathbf{r}' \mid \beta \mathbf{r}_m', \Sigma, \mathbf{I}) = \frac{\exp(-\frac{1}{2} \text{Tr}[(\mathbf{r}' - \alpha \mathbf{e}' - \beta \mathbf{r}_m')' \Sigma^{-1} (\mathbf{r}' - \alpha \mathbf{e}' - \beta \mathbf{r}_m')])}{(2\pi)^{nT/2} T^{n/2} |\Sigma|^{T/2}}$$

thus the log likelihood function is

$$\log L = -\frac{1}{2} \text{Tr}[(\mathbf{r}' - \alpha \mathbf{e}' - \beta \mathbf{r}_m')' \Sigma^{-1} (\mathbf{r}' - \alpha \mathbf{e}' - \beta \mathbf{r}_m')] - \frac{nT}{2} \log 2\pi - \frac{n}{2} \log T - \frac{T}{2} \log |\Sigma|$$

FOC w.r.t  $\alpha$ , by chain rule(Petersen and Pedersen 2012)

$$\begin{aligned} \partial \log L &= \text{Tr} \left( \frac{\partial \log L}{\partial (\mathbf{r}' - \alpha \mathbf{e}' - \beta \mathbf{r}_m')} \right)' \partial (\mathbf{X} - \alpha \mathbf{e}' - \beta \mathbf{r}_m') \\ &= \text{Tr} \left( \frac{\partial \log L}{\partial (\mathbf{r}' - \alpha \mathbf{e}' - \beta \mathbf{r}_m')} \right)' \partial \alpha (-\mathbf{e}') \\ &= \text{Tr} \left( (-\mathbf{e}') \frac{\partial \log L}{\partial (\mathbf{r}' - \alpha \mathbf{e}' - \beta \mathbf{r}_m')} \right)' \partial \alpha \end{aligned}$$

hence

$$\begin{aligned}\frac{\partial \log L}{\partial \alpha} &= -\frac{\partial \log L}{\partial (\mathbf{r}' - \alpha \mathbf{e}' - \beta \mathbf{r}'_{\mathbf{m}})} \mathbf{e} \\ &= -(\Sigma^- + \Sigma'^-)(\mathbf{r}' - \alpha \mathbf{e}' - \beta \mathbf{r}'_{\mathbf{m}}) \mathbf{e} = 0\end{aligned}$$

Similarly, FOC w.r.t  $\beta$  and combine those results:

$$\begin{aligned}(\mathbf{r}' - \hat{\alpha} \mathbf{e}' - \hat{\beta} \mathbf{r}'_{\mathbf{m}}) \mathbf{r}_{\mathbf{m}} &= \mathbf{0} \\ (\mathbf{r}' - \hat{\alpha} \mathbf{e}' - \hat{\beta} \mathbf{r}'_{\mathbf{m}}) \mathbf{e} &= \mathbf{0}\end{aligned}$$

which leads to a linear equation

$$\begin{bmatrix} \mathbf{e}' \mathbf{r}_{\mathbf{m}} & \mathbf{r}'_{\mathbf{m}} \mathbf{r}_{\mathbf{m}} \\ \mathbf{e}' \mathbf{e} & \mathbf{r}'_{\mathbf{m}} \mathbf{e} \end{bmatrix} \begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} = \begin{bmatrix} \mathbf{r}' \mathbf{r}_{\mathbf{m}} \\ \mathbf{r}' \mathbf{e} \end{bmatrix}$$

Similarly to our deduction for mean-variance model, let  $a = \mathbf{r}'_{\mathbf{m}} \mathbf{r}_{\mathbf{m}}$ ,  $b = \mathbf{e}' \mathbf{e} = T$  and  $c = \mathbf{e}' \mathbf{r}_{\mathbf{m}}$  ( $c^2 < ab$ ), hence

$$\begin{cases} \hat{\alpha} = c \mathbf{r}' \mathbf{r}_{\mathbf{m}} - a \mathbf{r}' \mathbf{e} / (c^2 - ab) \\ \hat{\beta} = -b \mathbf{r}' \mathbf{r}_{\mathbf{m}} + c \mathbf{r}' \mathbf{e} / (c^2 - ab) \end{cases}$$

By assumption  $\mathbf{r}' \sim \mathcal{MN}_{n \times T}(\beta \mathbf{r}'_{\mathbf{m}}, \Sigma, \mathbf{I})$ , and  $\hat{\alpha} = \mathbf{r}'(c \mathbf{r}_{\mathbf{m}} - a \mathbf{e}) / (c^2 - ab)$ . By transformation of matrix normal distribution (Wikipedia contributors 2019)

$$\frac{(c \mathbf{r}_{\mathbf{m}} - a \mathbf{e})'(c \mathbf{r}_{\mathbf{m}} - a \mathbf{e})}{(c^2 - ab)^2} = \frac{c^2 a - 2ac^2 + a^2 b}{(c^2 - ab)^2} = \frac{a}{ab - c^2}$$

we have

$$\hat{\alpha} \sim \mathcal{MN}(\mathbf{0}, \Sigma, \frac{a}{ab - c^2}) \sim \mathcal{N}(0, \frac{a}{ab - c^2} \Sigma)$$

which degenerated to multivariate normal distribution since  $\Sigma \otimes \frac{a}{ab - c^2} = \frac{a}{ab - c^2} \Sigma$ . For the same reason,  $\hat{\beta} \sim \mathcal{N}(\beta, \frac{b}{ab - c^2} \Sigma)$ .

Then we may construct statistic  $J_0$  as

$$J_0 = \hat{\alpha}' \left( \frac{a}{ab - c^2} \Sigma \right)^- \hat{\alpha} \sim \chi_n^2$$

However,  $\Sigma$  is unknown so we should use  $\hat{\Sigma}$  instead of  $\Sigma$  and now  $J_0$  is just asymptotically chi-square distributed:

$$J_0 = \frac{ab - c^2}{a} \hat{\alpha}' \hat{\Sigma}^- \hat{\alpha} \stackrel{A}{\sim} \chi_n^2$$

Return to the likelihood equation and FOC w.r.t  $\Sigma$ :

$$\frac{\partial \log L}{\partial \Sigma} = \frac{1}{2} (\Sigma^- (\mathbf{r}' - \alpha \mathbf{e}' - \beta \mathbf{r}'_{\mathbf{m}}) (\mathbf{r}' - \alpha \mathbf{e}' - \beta \mathbf{r}'_{\mathbf{m}})' \Sigma^-)' - \frac{T}{2} \Sigma'^- = 0$$

hence

$$\hat{\Sigma} = (\mathbf{r}' - \hat{\alpha} \mathbf{e}' - \hat{\beta} \mathbf{r}'_{\mathbf{m}}) (\mathbf{r}' - \hat{\alpha} \mathbf{e}' - \hat{\beta} \mathbf{r}'_{\mathbf{m}})' / T$$

Where

$$(\mathbf{r}' - \hat{\alpha}\mathbf{e}' - \hat{\beta}\mathbf{r}'_{\mathbf{m}}) = \mathbf{r}'(\mathbf{I} - \frac{c\mathbf{r}_{\mathbf{m}}\mathbf{e}' - a\mathbf{e}\mathbf{e}' - b\mathbf{r}_{\mathbf{m}}\mathbf{r}'_{\mathbf{m}} + c\mathbf{e}\mathbf{r}'_{\mathbf{m}}}{c^2 - ab})$$

Easy to verify  $\frac{c\mathbf{r}_{\mathbf{m}}\mathbf{e}' - a\mathbf{e}\mathbf{e}' - b\mathbf{r}_{\mathbf{m}}\mathbf{r}'_{\mathbf{m}} + c\mathbf{e}\mathbf{r}'_{\mathbf{m}}}{c^2 - ab}$  is symmetric and idempotent, thus

$$\begin{aligned} \text{rank}(\mathbf{I} - \frac{c\mathbf{r}_{\mathbf{m}}\mathbf{e}' - a\mathbf{e}\mathbf{e}' - b\mathbf{r}_{\mathbf{m}}\mathbf{r}'_{\mathbf{m}} + c\mathbf{e}\mathbf{r}'_{\mathbf{m}}}{c^2 - ab}) &= \text{Tr}(\mathbf{I} - \frac{c\mathbf{r}_{\mathbf{m}}\mathbf{e}' - a\mathbf{e}\mathbf{e}' - b\mathbf{r}_{\mathbf{m}}\mathbf{r}'_{\mathbf{m}} + c\mathbf{e}\mathbf{r}'_{\mathbf{m}}}{c^2 - ab}) \\ &= T - \frac{2c^2 - 2ab}{c^2 - ab} \\ &= T - 2 \end{aligned}$$

By following lemma:

**Lemma** Suppose symmetric matrix  $p \times p$   $\mathbf{A}$ . It's idempotent of rank  $s$  iff there exist a  $p \times s$   $\mathbf{P} \ni \mathbf{P}\mathbf{P}' = \mathbf{A}$  and  $\mathbf{P}'\mathbf{P} = \mathbf{I}$ .

**Proof** Sufficiency is trivial. For necessity, since  $\mathbf{A}$  is symmetric and idempotent matrix, it can be spectral decomposition by  $\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}'$ . Where the diagonal of  $\mathbf{\Lambda}$  is  $s$  1 and  $p - s$  0. Thus

$$\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}' = (\mathbf{P}_1 \quad \mathbf{P}_2) \begin{pmatrix} \mathbf{I}_s & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{P}'_1 \\ \mathbf{P}'_2 \end{pmatrix} = \mathbf{P}_1\mathbf{P}'_1$$

Note

$$\mathbf{I}_p = \mathbf{Q}'\mathbf{Q} = \begin{pmatrix} \mathbf{P}'_1 \\ \mathbf{P}'_2 \end{pmatrix} (\mathbf{P}_1 \quad \mathbf{P}_2) = \begin{pmatrix} \mathbf{P}'_1\mathbf{P}_1 & \mathbf{P}'_1\mathbf{P}_2 \\ \mathbf{P}'_2\mathbf{P}_1 & \mathbf{P}'_2\mathbf{P}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{P}'_1\mathbf{P}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{P}'_2\mathbf{P}_2 \end{pmatrix}$$

hence  $\mathbf{P}'_1\mathbf{P}_1 = \mathbf{I}_s$ . ■

We may find  $\mathbf{r}'\mathbf{P} \sim \mathcal{MN}_{n \times (T-2)}(\mathbf{0}, \Sigma, \mathbf{I})$ . Where  $\mathbb{E}[\mathbf{r}'\mathbf{P}] = \mathbb{E}[\mathbf{r}'\mathbf{P}\mathbf{P}'\mathbf{P}] = \mathbb{E}[\mathbf{r}'\mathbf{A}\mathbf{P}] = \mathbf{0}$  and thus (Wikipedia contributors 2020)

$$T\hat{\Sigma} = \mathbf{r}'\mathbf{A}\mathbf{r} = \mathbf{r}'\mathbf{P}\mathbf{P}'\mathbf{r} = \mathbf{r}'\mathbf{P}(\mathbf{r}'\mathbf{P})' \sim W_n(T-2, \Sigma)$$

**Theorem** Suppose  $\mathbf{A} \sim W_n(m, \Sigma)$  and  $\mathbf{x} \sim \mathcal{N}(\mathbf{0}, \Sigma)$ , where  $m > n$ , then

$$\frac{m-n+1}{n} \mathbf{x}'\mathbf{A}^{-}\mathbf{x} \sim F_{n, m-n+1}$$

**Proof** Note

$$\mathbf{x}'\mathbf{A}^{-}\mathbf{x} = \frac{\mathbf{x}'\Sigma^{-}\mathbf{x}}{\mathbf{x}'\Sigma^{-}\mathbf{x} / \mathbf{x}'\mathbf{A}^{-}\mathbf{x}}$$

where  $\mathbf{x}'\Sigma^{-}\mathbf{x} \sim \chi_n^2$  and  $\frac{\mathbf{x}'\Sigma^{-}\mathbf{x}}{\mathbf{x}'\mathbf{A}^{-}\mathbf{x}} \sim \chi_{m-n+1}^2$  [gupta2018matrix]. Then this claim follows. ■

Thus, taking  $\mathbf{x} = \sqrt{\frac{ab-c^2}{a}}\hat{\alpha}$  and apply above theorem, we have

$$J_1 = \frac{T-n-1}{n} \frac{ab-c^2}{a} \frac{1}{T} \hat{\alpha}'\hat{\Sigma}^{-}\hat{\alpha} = \frac{T-n-1}{nT} J_0 \sim F_{n, T-n-1}$$

$J_1$  is the so called GRS statistic.

### 0.1 Interpretation of $J_1$

Consider a sub portfolio problem:

$$\min \frac{1}{2} \omega' \mathbf{V} \omega \quad s.t. \quad \mathbf{e}' \omega - \mathbf{1}, \omega' \bar{\mathbf{r}} = m$$

where  $r = \begin{bmatrix} r_{m,t}^e \\ \mathbf{r}_t^e \end{bmatrix}$ . Thus the MLE of  $\bar{\mathbf{r}}$

$$\hat{\mathbf{r}} = \mathbb{E} \left[ \widehat{\begin{bmatrix} r_{m,t}^e \\ \mathbf{r}_t^e \end{bmatrix}} \right] = \frac{1}{T} \begin{bmatrix} \mathbf{r}_m' \mathbf{e} \\ \mathbf{r}' \mathbf{e} \end{bmatrix}$$

and  $\mathbf{V}$  follows by  $\hat{\sigma}^2 = \mathbf{r}_m' (\mathbf{I} - \mathbf{P}_1) \mathbf{r}_m = a - \frac{c^2}{b}$  and recall

$$\mathbf{r}_t^e = \alpha + \beta r_{m,t}^e + \nu_t$$

then

$$\text{Cov}(\mathbf{r}_t^e, r_{m,t}^e) = \text{Cov}(\beta r_{m,t}^e, r_{m,t}^e) = \beta \hat{\sigma}^2$$

$$\text{Cov}(\mathbf{r}_t^e) = \text{Cov}(\beta r_{m,t}^e) + \text{Cov}(\nu_t) = \beta \sigma^2 \beta' + \Sigma$$

hence

$$\hat{\mathbf{V}} = \text{Cov}(\bar{\mathbf{r}}) = \begin{bmatrix} \hat{\sigma}^2 & \hat{\sigma}^2 \beta' \\ \hat{\sigma}^2 \beta & \Sigma + \beta \hat{\sigma}^2 \beta' \end{bmatrix}$$

Using the formula for a partitioned inverse:

$$\begin{aligned} \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} &= \begin{bmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1} \mathbf{B} (\mathbf{D} - \mathbf{C} \mathbf{A}^{-1} \mathbf{B})^{-1} \mathbf{C} \mathbf{A}^{-1} & -\mathbf{A}^{-1} \mathbf{B} (\mathbf{D} - \mathbf{C} \mathbf{A}^{-1} \mathbf{B})^{-1} \\ -(\mathbf{D} - \mathbf{C} \mathbf{A}^{-1} \mathbf{B})^{-1} \mathbf{C} \mathbf{A}^{-1} & (\mathbf{D} - \mathbf{C} \mathbf{A}^{-1} \mathbf{B})^{-1} \end{bmatrix} \\ &= \begin{bmatrix} (\mathbf{A} - \mathbf{B} \mathbf{D}^{-1} \mathbf{C})^{-1} & -(\mathbf{A} - \mathbf{B} \mathbf{D}^{-1} \mathbf{C})^{-1} \mathbf{B} \mathbf{D}^{-1} \\ -\mathbf{D}^{-1} \mathbf{C} (\mathbf{A} - \mathbf{B} \mathbf{D}^{-1} \mathbf{C})^{-1} & \mathbf{D}^{-1} + \mathbf{D}^{-1} \mathbf{C} (\mathbf{A} - \mathbf{B} \mathbf{D}^{-1} \mathbf{C})^{-1} \mathbf{B} \mathbf{D}^{-1} \end{bmatrix} \end{aligned}$$

Thus

$$\hat{\mathbf{V}}^{-1} = \begin{bmatrix} \frac{1}{\hat{\sigma}^2} + \beta' \Sigma^{-1} \beta & -\beta' \Sigma^{-1} \\ -\Sigma^{-1} \beta & \Sigma^{-1} \end{bmatrix}$$

Note the maximal sharpe ratio is given by

$$S_p = \sqrt{\bar{\mathbf{r}}' \hat{\mathbf{V}}^{-1} \bar{\mathbf{r}}}$$

Directly algebra and the invariance property of MLE leads to

$$\hat{\mathbf{r}}' \hat{\mathbf{V}}^{-1} \hat{\mathbf{r}} = \frac{\mathbf{r}_m' \mathbf{e}^2}{T^2 \hat{\sigma}^2} + \frac{(\mathbf{r}' \mathbf{e} - \hat{\beta} \mathbf{r}_m' \mathbf{e})' \hat{\Sigma}^{-1} (\mathbf{r}' \mathbf{e} - \hat{\beta} \mathbf{r}_m' \mathbf{e})}{T^2}$$

note  $(\mathbf{r}' - \hat{\alpha} \mathbf{e}' - \hat{\beta} \mathbf{r}_m' \mathbf{e}) \mathbf{e} = \mathbf{0}$ :

$$\hat{\alpha} \hat{\Sigma}^{-1} \hat{\alpha} = \hat{\mathbf{r}}' \hat{\mathbf{V}}^{-1} \bar{\mathbf{r}} - \frac{\mathbf{r}_m' \mathbf{P}_1 \mathbf{r}_m}{\mathbf{r}_m' (\mathbf{I} - \mathbf{P}_1) \mathbf{r}_m}$$

where  $\frac{\mathbf{r}'_m \mathbf{P}_1 \mathbf{r}_m}{\mathbf{r}'_m (\mathbf{I} - \mathbf{P}_1) \mathbf{r}_m}$  is the MLE of market Sharpe ratio.

Note  $ab - c^2$  can be seen as estimator of variance of  $r_{m,t}$  since

$$\hat{\theta}_m^2 = \frac{\mathbf{r}'_m \mathbf{P}_1 \mathbf{r}_m}{\mathbf{r}'_m (\mathbf{I} - \mathbf{P}_1) \mathbf{r}_m} = \frac{\frac{c^2}{b}}{a - \frac{c^2}{b}} = \frac{c^2}{ab - c^2}$$

That is

$$J_1 = \frac{T - n - 1}{n} \frac{\hat{\theta}_p^2 - \hat{\theta}_m^2}{1 + \hat{\theta}_m^2}$$

- Petersen, K. B., and M. S. Pedersen. 2012. “The Matrix Cookbook.” Technical University of Denmark. <http://www2.compute.dtu.dk/pubdb/pubs/3274-full.html>.
- Wikipedia contributors. 2019. “Matrix Normal Distribution — Wikipedia, the Free Encyclopedia.” [https://en.wikipedia.org/w/index.php?title=Matrix\\_normal\\_distribution&oldid=902125596](https://en.wikipedia.org/w/index.php?title=Matrix_normal_distribution&oldid=902125596).
- . 2020. “Wishart Distribution — Wikipedia, the Free Encyclopedia.” [https://en.wikipedia.org/w/index.php?title=Wishart\\_distribution&oldid=986003757](https://en.wikipedia.org/w/index.php?title=Wishart_distribution&oldid=986003757).