MATRIX APPROCH TO GRS STATISTIC

Xie zejian

xiezej@gmail.com

December 14, 2020

1 GRS for CAPM

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

$$\mathbf{r_t^e} = \alpha + \beta r_{m,t}^e + \epsilon_\mathbf{t}$$

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, $\boldsymbol{\alpha} = \mathbf{0}$ 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

$$\mathbf{r} = \begin{bmatrix} \mathbf{r_1^e}' \\ \mathbf{r_2^e}' \\ \mathbf{r_3^e}' \\ \vdots \\ \mathbf{r_T^e}' \end{bmatrix}, \mathbf{r_m} = \begin{bmatrix} r_{m,1}^e \\ r_{m,2}^e \\ r_{m,3}^e \\ \vdots \\ r_{m,T}^e \end{bmatrix}$$

The equation become

$$\mathbf{r}' = \boldsymbol{\alpha}\mathbf{e}' - \boldsymbol{\beta}\mathbf{r}_\mathbf{m}' + \mathbf{E}$$

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

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

thus the log likelihood function is

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

FOC w.r.t α, by chain rule(Petersen and Pedersen 2012)

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

hence

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

Similarly, FOC w.r.t β 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

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

Similarly to our deduction for mean-variance mdoel, 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

$$
\hat{\alpha} = \frac{\mathbf{r}'(\mathbf{I} - \mathbf{P}_{\mathbf{r}_\mathbf{m}})\mathbf{e}}{\mathbf{e}'(\mathbf{I} - \mathbf{P}_{\mathbf{r}_\mathbf{m}})\mathbf{e}}
$$

and note

$$
\hat{\boldsymbol{\beta}}\mathbf{r}_\mathbf{m}' = \mathbf{r}' \frac{\mathbf{e}'(\mathbf{I} - \mathbf{P}_{\mathbf{r}_\mathbf{m}})\mathbf{e}\mathbf{P}_{\mathbf{r}_\mathbf{m}} - (\mathbf{I} - \mathbf{P}_{\mathbf{r}_\mathbf{m}})\mathbf{e}\mathbf{e}'\mathbf{P}_{\mathbf{r}_\mathbf{m}}}{\mathbf{e}'(\mathbf{I} - \mathbf{P}_{\mathbf{r}_\mathbf{m}})\mathbf{e}}
$$

By assumption $\mathbf{r}' \sim \mathcal{MN}_{n\times T}(\boldsymbol{\beta}\mathbf{r}_\mathbf{m}', \mathbf{\Sigma}, \mathbf{I})$. The transformation of matrix normal distribution(Wikipedia contributors 2019) yields

$$
\frac{[(\mathbf{I} - \mathbf{P}_{\mathbf{r}_\mathbf{m}})\mathbf{e}]'[(\mathbf{I} - \mathbf{P}_{\mathbf{r}_\mathbf{m}})\mathbf{e}]}{[\mathbf{e}'(\mathbf{I} - \mathbf{P}_{\mathbf{r}_\mathbf{m}})\mathbf{e}]^2} = \frac{\mathbf{1}}{\mathbf{e}'(\mathbf{I} - \mathbf{P}_{\mathbf{r}_\mathbf{m}})\mathbf{e}}
$$

we have

$$
\hat{\alpha} \sim \mathcal{MN}(\mathbf{0}, \mathbf{\Sigma}, \frac{\mathbf{1}}{\mathbf{e}'(\mathbf{I} - \mathbf{P}_{\mathbf{r}_\mathbf{m}})\mathbf{e}}) \sim \mathcal{N}(0, \frac{\mathbf{1}}{\mathbf{e}'(\mathbf{I} - \mathbf{P}_{\mathbf{r}_\mathbf{m}})\mathbf{e}}\mathbf{\Sigma})
$$

which degenerated to mutivariate normal distribution since $\mathbf{\Sigma} \otimes \frac{\mathbf{1}}{\mathbf{e}'(\mathbf{I} - \mathbf{P}_{\mathbf{r}_\mathbf{m}})\mathbf{e}} = \frac{\mathbf{1}}{\mathbf{e}'(\mathbf{I} - \mathbf{P}_{\mathbf{r}_\mathbf{m}})\mathbf{e}}\mathbf{\Sigma}$.

Then we may construct statistic J_0 as

$$
J_0 = \hat{\boldsymbol{\alpha}}'(\frac{\mathbf{1}}{\mathbf{e}'(\mathbf{I} - \mathbf{P}_{\mathbf{r}_\mathbf{m}})\mathbf{e}}\mathbf{\Sigma})^- \hat{\boldsymbol{\alpha}} \sim \chi_n^2
$$

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

$$J_0 = \mathbf{e'(I - P_{r_m})e}\hat{\boldsymbol{\alpha}}'\hat{\boldsymbol{\Sigma}}^{-}\hat{\boldsymbol{\alpha}} \overset{A}{\sim} \chi^2_n$$

Return to the likelihood equation and FOC w.r.t Σ:

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

hence

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

note

$$\hat{\boldsymbol{\alpha}}\mathbf{e}' + \hat{\boldsymbol{\beta}}\mathbf{r}_{\mathbf{m}}' = \mathbf{r}'\frac{\mathbf{e'(I - P_{r_m})eP_{r_m} + (I - P_{r_m})ee'(I - P_{r_m})}}{\mathbf{e'(I - P_{r_m})e}}$$

Thus

$$(\mathbf{r}' - \hat{\alpha}\mathbf{e}' - \hat{\beta}\mathbf{r}_{\mathbf{m}}') = \mathbf{r}'\left(\mathbf{I} - \frac{\mathbf{e'(I - P_{r_m})eP_{r_m} + (I - P_{r_m})ee'(I - P_{r_m})}}{\mathbf{e'(I - P_{r_m})e}}\right)$$

The transform matrix is symmetric and idempotent, thus

$$\begin{aligned}\text{rank}(\mathbf{I} - \frac{\mathbf{e'(I - P_{r_m})eP_{r_m} + (I - P_{r_m})ee'(I - P_{r_m})}}{\mathbf{e'(I - P_{r_m})e}}) &= \text{Tr}(\mathbf{I} - \frac{\mathbf{e'(I - P_{r_m})eP_{r_m} + (I - P_{r_m})ee'(I - P_{r_m})}}{\mathbf{e'(I - P_{r_m})e}}) \\ &= T - (1+1) \\ &= T - 2\end{aligned}$$

By following lemma:

> **Lemma** Suppose symmetric matrix $p \times p$ $\mathbf{A}$. It's idempotent of rank s iff there exist a $p \times s$ $\mathbf{P}$ $\ni \mathbf{PP' = A}$ and $\mathbf{P'P = I}$.
> **Proof** Sufficiency is trivial. For necessity, since $\mathbf{A}$ is symmetric and idempotent matrix, it can be spectral decompostioned by $\mathbf{A = Q\Lambda Q'}$. Where the diagonal of $\boldsymbol{\Lambda}$ is s 1 and $p-s$ 0. Thus
>
> $$\mathbf{A = Q\Lambda Q'} = \begin{pmatrix} \mathbf{P_1} & \mathbf{P_2} \end{pmatrix}\begin{pmatrix} \mathbf{I}_s & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \mathbf{P}_1' \\ \mathbf{P}_2' \end{pmatrix} = \mathbf{P}_1\mathbf{P}_1'$$
>
> Note
>
> $$\mathbf{I}_p = \mathbf{Q}'\mathbf{Q} = \begin{pmatrix} \mathbf{P}_1' \\ \mathbf{P}_2' \end{pmatrix}\begin{pmatrix} \mathbf{P}_1 & \mathbf{P}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{P}_1'\mathbf{P}_1 & \mathbf{P}_1'\mathbf{P}_2 \\ \mathbf{P}_2'\mathbf{P}_1 & \mathbf{P}_2'\mathbf{P}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{P}_1'\mathbf{P}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_2'\mathbf{P}_2 \end{pmatrix}$$
>
> hence$\mathbf{P_1'P_1 = I_s}$.■

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

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

> **Theorem** Suppose $\mathbf{A} \sim W_n(m, \Sigma)$ and $\mathbf{x} \sim \mathcal{N}_\setminus(\mathbf{0}, \Sigma)$, where $m > n$, then
>
> $$\frac{m-n+1}{n}\mathbf{x'A^{-}x} \sim F_{n,m-n+1}$$
>
> **Proof** Note
>
> $$\mathbf{x'A^{-}x} = \frac{\mathbf{x'\Sigma^{-}x}}{\mathbf{x'\Sigma^{-}x'/x'A^{-}x}}$$
>
> where $\mathbf{x'\Sigma^{-}x} \sim \chi^2_n$ and $\frac{\mathbf{x'\Sigma^{-}x}}{\mathbf{x'A^{-}x}} \sim \chi^2_{m-n+1}$(Gupta and Nagar 2018). Then this claim follows. ■

Thus, taking $\mathbf{x} = \sqrt{\mathbf{e}'(\mathbf{I}-\mathbf{P}_{\mathbf{r_m}})\mathbf{e}}\hat{\boldsymbol{\alpha}}$ and apply above theorem, we have

$$J_1 = \frac{T-n-1}{n}\frac{1}{T}\mathbf{e}'(\mathbf{I}-\mathbf{P}_{\mathbf{r_m}})\mathbf{e}\hat{\boldsymbol{\alpha}}'\hat{\boldsymbol{\Sigma}}^{-}\hat{\boldsymbol{\alpha}} = \frac{T-n-1}{nT}J_0 \sim F_{n,T-n-1}$$

J_1 is the so called GRS statistic.

1.1 Interpretion of J_1

Consider a sub portfolio problem:

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

where $r = \begin{bmatrix} r^e_{m,t} \\ \mathbf{r^e_t} \end{bmatrix}$. Thus the MLE of $\overline{\mathbf{r}}$

$$\hat{\mathbf{r}} = \mathrm{E}[\widehat{\begin{bmatrix} r^e_{m,t} \\ \mathbf{r^e_t} \end{bmatrix}}] = \frac{1}{T}\begin{bmatrix} \mathbf{r'_m e} \\ \mathbf{r'e} \end{bmatrix}$$

and $\mathbf{V}$ comes from $\hat{\sigma}^2 = \frac{\mathbf{r'_m(I-P_1)r_m}}{\mathbf{e'e}}$ and recall

$$\mathbf{r^e_t} = \boldsymbol{\alpha} + \beta r^e_{m,t} + \nu_t$$

then

$$\mathrm{Cov}(\mathbf{r^e_t}, r^e_{m,t}) = \mathrm{Cov}(\beta r^e_{m,t}, r^e_{m,t}) = \beta\hat{\sigma}^2$$

$$\mathrm{Cov}(\mathbf{r^e_t}) = \mathrm{Cov}(\beta r^e_{m,t}) + \mathrm{Cov}(\nu_t) = \beta\sigma^2\beta' + \Sigma$$

hence

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

Using the formula for a partitioned inverse:

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

Thus

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

Note the maximal sharpe ratio is given by

$$S_p = \sqrt{\overline{\mathbf{r}}'\mathbf{V}^{-}\overline{\mathbf{r}}}$$

Directly algebra and the invariance property of MLE yields

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

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

$$\hat{\alpha}\hat{\mathbf{\Sigma}}^{-}\hat{\alpha} = \hat{\mathbf{r}}'\hat{\mathbf{V}}^{-}\overline{\mathbf{r}} - \frac{\mathbf{r}_\mathbf{m}'\mathbf{P_1}\mathbf{r_m}}{\mathbf{r}_\mathbf{m}'(\mathbf{I}-\mathbf{P_1})\mathbf{r_m}}$$

where $\frac{\mathbf{r}_\mathbf{m}'\mathbf{P_1}\mathbf{r_m}}{\mathbf{r}_\mathbf{m}'(\mathbf{I}-\mathbf{P_1})\mathbf{r_m}}$ is the MLE of square market Sharpe ratio. Together with

$$1 + \hat{\theta_m}^2 = 1 + \frac{\mathbf{r}_\mathbf{m}'\mathbf{P_1}\mathbf{r_m}}{\mathbf{r}_\mathbf{m}'(\mathbf{I}-\mathbf{P_1})\mathbf{r_m}} = \frac{\mathbf{r}_\mathbf{m}'\mathbf{r_m}}{\mathbf{r}_\mathbf{m}'(\mathbf{I}-\mathbf{P_1})\mathbf{r_m}} = \frac{\mathbf{e}'\mathbf{e}}{\mathbf{e}'(\mathbf{I}-\mathbf{P}_{\mathbf{r_m}})\mathbf{e}}$$

we finally have:

$$J_1 = \frac{T-n-1}{n}\frac{\hat{\theta_p}^2 - \hat{\theta_m}^2}{1+\hat{\theta_m}^2}$$

2 GRS for multi-factor model

Follow the similar fushion:

$$\mathbf{r}' = \boldsymbol{\alpha}\mathbf{e}' + \mathbf{B}\mathbf{F}' + \mathbf{E}$$

where $\mathbf{B}$ is $n \times k$ and $\mathbf{F}'$ is $k \times t$

Now $\mathbf{r}' \sim \mathcal{MN}_{n\times T}(\boldsymbol{\alpha}\mathbf{e}' + \mathbf{B}\mathbf{F}', \mathbf{\Sigma}, \mathbf{I})$, the MLE satisfy:

$$(\mathbf{r}' - \hat{\boldsymbol{\alpha}}\mathbf{e}' - \hat{\mathbf{B}}\mathbf{F}')\mathbf{F} = \mathbf{0}$$
$$(\mathbf{r}' - \hat{\boldsymbol{\alpha}}\mathbf{e}' - \hat{\mathbf{B}}\mathbf{F}')\mathbf{e} = \mathbf{0}$$

which leads to a linear equation

$$\begin{bmatrix}\mathbf{F}'\mathbf{F} & \mathbf{F}'\mathbf{e}\\ \mathbf{e}'\mathbf{F} & \mathbf{e}'\mathbf{e}\end{bmatrix}\begin{bmatrix}\hat{\mathbf{B}}'\\ \hat{\boldsymbol{\alpha}}'\end{bmatrix} = \begin{bmatrix}\mathbf{F}'\mathbf{r}\\ \mathbf{e}'\mathbf{r}\end{bmatrix}$$

Recall the inverse of partitioned matrix:

$$\begin{aligned}\begin{bmatrix}\hat{\mathbf{B}}'\\ \hat{\boldsymbol{\alpha}}'\end{bmatrix} &= \begin{bmatrix}\mathbf{F}'\mathbf{F} & \mathbf{F}'\mathbf{e}\\ \mathbf{e}'\mathbf{F} & \mathbf{e}'\mathbf{e}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{F}'\mathbf{r}\\ \mathbf{e}'\mathbf{r}\end{bmatrix}\\ &= \frac{1}{\mathbf{e}'(\mathbf{I}-\mathbf{P_F})\mathbf{e}}\begin{bmatrix}\mathbf{e}'(\mathbf{I}-\mathbf{P_F})\mathbf{e}(\mathbf{F}'\mathbf{F})^{-1} + (\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'\mathbf{e}\mathbf{e}'\mathbf{F}(\mathbf{F}'\mathbf{F})^{-1} & -(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'\mathbf{e}\\ -\mathbf{e}'\mathbf{F}(\mathbf{F}'\mathbf{F})^{-1} & 1\end{bmatrix}\begin{bmatrix}\mathbf{F}'\mathbf{r}\\ \mathbf{e}'\mathbf{r}\end{bmatrix}\end{aligned}$$

Thus

$$\hat{\boldsymbol{\alpha}} = \frac{\mathbf{r}'(\mathbf{I}-\mathbf{P_F})\mathbf{e}}{\mathbf{e}'(\mathbf{I}-\mathbf{P_F})\mathbf{e}} \sim \mathcal{N}(0, \frac{\mathbf{\Sigma}}{\mathbf{e}'(\mathbf{I}-\mathbf{P_F})\mathbf{e}})$$

and note

$$\hat{\boldsymbol{\alpha}}\mathbf{e}' + \hat{\mathbf{B}}\mathbf{F}' = \mathbf{r}'\frac{(\mathbf{I}-\mathbf{P_F})\mathbf{e}\mathbf{e}' + \mathbf{e}'(\mathbf{I}-\mathbf{P_F})\mathbf{e}\mathbf{P_F} - (\mathbf{I}-\mathbf{P_F})\mathbf{e}\mathbf{e}'\mathbf{P_F}}{\mathbf{e}'(\mathbf{I}-\mathbf{P_F})\mathbf{e}}$$
$$= \mathbf{r}'\frac{(\mathbf{I}-\mathbf{P_F})\mathbf{e}\mathbf{e}'(\mathbf{I}-\mathbf{P_F}) + \mathbf{e}'(\mathbf{I}-\mathbf{P_F})\mathbf{e}\mathbf{P_F}}{\mathbf{e}'(\mathbf{I}-\mathbf{P_F})\mathbf{e}}$$

Return to the likelihood equation and FOC w.r.t Σ:

$$\hat{\Sigma} = (\mathbf{r}' - \hat{\alpha}\mathbf{e}' - \hat{\mathbf{B}}\mathbf{F}')(\mathbf{r}' - \hat{\alpha}\mathbf{e}' - \hat{\mathbf{B}}\mathbf{F}')'/T$$

plug in $\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}}$:

$$(\mathbf{r}' - \hat{\boldsymbol{\alpha}}\mathbf{e}' - \hat{\mathbf{B}}\mathbf{F}') = \mathbf{r}'\left(\mathbf{I} - \frac{\mathbf{e}'(\mathbf{I}-\mathbf{P_F})\mathbf{e}\mathbf{P_F} + (\mathbf{I}-\mathbf{P_F})\mathbf{e}\mathbf{e}'(\mathbf{I}-\mathbf{P_F})}{\mathbf{e}'(\mathbf{I}-\mathbf{P_F})\mathbf{e}}\right)$$

$$\operatorname{rank}(\mathbf{I} - \frac{\mathbf{e}'(\mathbf{I}-\mathbf{P_F})\mathbf{e}\mathbf{P_F} + (\mathbf{I}-\mathbf{P_F})\mathbf{e}\mathbf{e}'(\mathbf{I}-\mathbf{P_F})}{\mathbf{e}'(\mathbf{I}-\mathbf{P_F})\mathbf{e}}) = \operatorname{Tr}(\mathbf{I} - \frac{\mathbf{e}'(\mathbf{I}-\mathbf{P_F})\mathbf{e}\mathbf{P_F} + (\mathbf{I}-\mathbf{P_F})\mathbf{e}\mathbf{e}'(\mathbf{I}-\mathbf{P_F})}{\mathbf{e}'(\mathbf{I}-\mathbf{P_F})\mathbf{e}})$$
$$= T - (k+1)$$
$$= T - k - 1$$

thus

$$T\hat{\Sigma} \sim W_n(T-k-1, \Sigma)$$

it's followed by

$$J_1 = \frac{T-n-k}{n}\frac{1}{T}\mathbf{e}'(\mathbf{I}-\mathbf{P_F})\mathbf{e}\hat{\boldsymbol{\alpha}}'\hat{\boldsymbol{\Sigma}}^{-}\hat{\boldsymbol{\alpha}} \sim F_{n,T-n-k}$$

Then we show that

$$\frac{\mathbf{e}'\mathbf{e}}{\mathbf{e}'(\mathbf{I}-\mathbf{P_F})\mathbf{e}} = \mathbf{1} + \hat{\mathbf{f}}'\widehat{\operatorname{Cov}(\mathbf{f})}^{-1}\hat{\mathbf{f}}$$

Where

$$\widehat{\operatorname{Cov}(\mathbf{f})} = \frac{\mathbf{F}'(\mathbf{I}-\mathbf{P_1})\mathbf{F}}{\mathbf{e}'\mathbf{e}}, \hat{\mathbf{f}} = \frac{\mathbf{F}'\mathbf{e}}{\mathbf{e}'\mathbf{e}}$$

It's sufficient to show that

$$\mathbf{F}[\mathbf{F}'(\mathbf{I}-\mathbf{P_1})\mathbf{F}]^{-1}\mathbf{F}'\mathbf{e} = \frac{\mathbf{P_F}\mathbf{e}\mathbf{e}'\mathbf{e}}{\mathbf{e}'(\mathbf{I}-\mathbf{P_F})\mathbf{e}}$$

Lemma Given any 2 matrix A and B, we claim that

$$\mathbf{P_A}\mathbf{B}[\mathbf{B}'(\mathbf{I}-\mathbf{P_A})\mathbf{B}]^{-1}\mathbf{B}'\mathbf{B} = \mathbf{A}[\mathbf{A}'(\mathbf{I}-\mathbf{P_B})\mathbf{A}]^{-1}\mathbf{A}'\mathbf{B}$$

Proof

$$\begin{aligned}
&\mathbf{P_A}\mathbf{B}[\mathbf{B}'(\mathbf{I}-\mathbf{P_A})\mathbf{B}]^{-1}\mathbf{B}'\mathbf{B} = \mathbf{A}[\mathbf{A}'(\mathbf{I}-\mathbf{P_B})\mathbf{A}]^{-1}\mathbf{A}'\mathbf{B} \\
\Longleftarrow\ &(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{B}[\mathbf{B}'(\mathbf{I}-\mathbf{P_A})\mathbf{B}]^{-1}\mathbf{B}'\mathbf{B} = [\mathbf{A}'(\mathbf{I}-\mathbf{P_B})\mathbf{A}]^{-1}\mathbf{A}'\mathbf{B} \\
\Longleftarrow\ &\mathbf{A}'(\mathbf{I}-\mathbf{P_B})\mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{B}[\mathbf{B}'(\mathbf{I}-\mathbf{P_A})\mathbf{B}]^{-1}\mathbf{B}'\mathbf{B} = \mathbf{A}'\mathbf{B} \\
\Longleftarrow\ &\mathbf{A}'(\mathbf{I}-\mathbf{P_B})\mathbf{P_A}\mathbf{B}[\mathbf{B}'(\mathbf{I}-\mathbf{P_A})\mathbf{B}]^{-1}\mathbf{B}'\mathbf{B} = \mathbf{A}'\mathbf{B} \\
\Longleftarrow\ &\mathbf{A}'(\mathbf{I}-\mathbf{P_B})\mathbf{P_A}\mathbf{B}[\mathbf{B}'(\mathbf{I}-\mathbf{P_A})\mathbf{B}]^{-1} = \mathbf{A}'\mathbf{B}[\mathbf{B}'\mathbf{B}]^{-1} \\
\Longleftarrow\ &\mathbf{A}'(\mathbf{I}-\mathbf{P_B})\mathbf{P_A}\mathbf{B} = \mathbf{A}'\mathbf{B}[\mathbf{B}'\mathbf{B}]^{-1}\mathbf{B}'(\mathbf{I}-\mathbf{P_A})\mathbf{B} \\
\Longleftarrow\ &\mathbf{A}'(\mathbf{I}-\mathbf{P_B})\mathbf{P_A}\mathbf{B} = \mathbf{A}'\mathbf{P_B}(\mathbf{I}-\mathbf{P_A})\mathbf{B}
\end{aligned}$$

The claim follows from both sides equal to $\mathbf{A'(I - P_B P_A)B}$. ■

Thus,

$$J_1 = \frac{T-n-k}{n} \frac{\hat{\boldsymbol{\alpha}}'\hat{\boldsymbol{\Sigma}}^{-}\hat{\boldsymbol{\alpha}}}{1+\hat{\theta_m}^2} \sim F_{n,T-n-k}$$

where

$$\hat{\theta_m}^2 = \hat{\mathbf{f}}'\widehat{\mathrm{Cov}(\mathbf{f})}^{-1}\hat{\mathbf{f}} = \frac{\mathbf{e'F[F'(I-P_1)F]^{-1}F'e}}{\mathbf{e'e}}$$

denoted the MLE of the maximal squared sharpe ration generated by the k risk factor $\mathbf{F'}$.

2.1 Interpretion of J_1

Consider a sub portfolio problem:

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

where $\mathbf{r} = \begin{bmatrix} \mathbf{f_t} \\ \mathbf{r_t^e} \end{bmatrix}$. Thus the MLE of $\overline{\mathbf{r}}$

$$\hat{\mathbf{r}} = \frac{1}{T}\begin{bmatrix} \mathbf{F'e} \\ \mathbf{r'e} \end{bmatrix}$$

and $\mathbf{V}$ is given by

$$\hat{\mathbf{V}} = \begin{bmatrix} \frac{\mathbf{F'(I-P_1)F}}{\mathbf{e'e}} & \frac{\mathbf{F'(I-P_1)F}}{\mathbf{e'e}}\mathbf{B'} \\ \mathbf{B}\frac{\mathbf{F'(I-P_1)F}}{\mathbf{e'e}} & \hat{\boldsymbol{\Sigma}} + \mathbf{B}\frac{\mathbf{F'(I-P_1)F}}{\mathbf{e'e}}\mathbf{B'} \end{bmatrix}$$

Directly algebra and the invariance property of MLE yields

$$\hat{\theta_p}^2 = \hat{\mathbf{r}}'\hat{\mathbf{V}}^{-}\hat{\mathbf{r}} = \frac{\mathbf{e'F}[\frac{\mathbf{F'(I-P_1)F}}{\mathbf{e'e}}]^{-1}\mathbf{F'e}}{T^2} + \frac{(\mathbf{r'e} - \hat{\mathbf{B}}\mathbf{F'e})'\hat{\boldsymbol{\Sigma}}^{-}(\mathbf{r'e} - \hat{\mathbf{B}}\mathbf{F'e})}{T^2}$$

that is

$$\hat{\alpha}\hat{\boldsymbol{\Sigma}}^{-}\hat{\alpha} = \hat{\theta_p}^2 - \hat{\theta_m}^2$$

we finally have:

$$J_1 = \frac{T-n-k}{n}\frac{\hat{\theta_p}^2 - \hat{\theta_m}^2}{1+\hat{\theta_m}^2}$$

Gupta, Arjun K, and Daya K Nagar. 2018. *Matrix Variate Distributions*. Vol. 104. CRC Press.

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`.

———. 2020. "Wishart Distribution — Wikipedia, the Free Encyclopedia." `https://en.wikipedia.org/w/index.php?title=Wishart_distribution&oldid=986003757`.