『学習院大学　経済論集』第38巻　第1号（2001年4月）

二次計画問題における最尤推定量の漸近的非許容性

Asymptotic Inadomissibility of Maximum Likelihood Estimator in
A Quadratic Programming Problem

森　治憲

Harunori Mori

1 はじめに

数理計画法が適用されるとき，その係数には推定値を用いることが多い。係数を推定値で代用した場合，推定誤差の影響は最適化へ波及する。それならば，どのような推定量が望ましいのだろうか。

本研究では多変量正規分布$N_K(\mu, \Sigma)$の母数を係数とする二次計画問題

$$\max\ a\boldsymbol{\mu}'\mathbf{x} - \mathbf{x}'\Sigma\mathbf{x}$$
$$s.t.\ \mathbf{p}'\mathbf{x} = M\ ,\ \mathbf{x} \geq \mathbf{0}$$

と最尤推定量との関係を考察した。ただし，$\mathbf{p}$はK次元列ベクトルである。そして，推定の評価基準を目的関数に関する平均二乗誤差としたとき，次の条件

$$a \geq M\Big/\sqrt{\mathbf{p}'\Sigma^{-1}\mathbf{p}}$$

が成立するならば，平均ベクトル$\boldsymbol{\mu}$と分散共分散行列Σの最尤推定量が漸近的に非許容的であることを証明した。さらに，最尤推定量を漸近的に優越する平均ベクトルの推定量が縮小推定量であることも示すことができた。

$\mathbf{p}=\mathbf{1}$, $M=1$とすると，この二次計画問題は資産選択問題における平均・分散モデルとなる[1]。そのため，最適化に対する推定誤差の影響は金融工学の分野で関心が高く，多くの研究がこの問題を取り上げている。しかし，陽表的に表すことができない二次計画問題の最適解の標本分布を求めることは非常に難しく，統計学的に十分な成果が得られているとは言い難い。

ここで最適解（最適資産配分量）の統計的性質に関連した，これまでの研究を簡単に紹介しておく。Jobson and Korkie［3］は平均と分散を不偏推定量で代用したときの最適解の漸近分布を導出した。最適解の統計的性質を初めて明らかにした点は評価できるが，この研究は具体的に推定量を提案したものではない。最尤推定量の下で近似的な最適解の標本分布を求めたBritten-Jones［1］についても同様である。平均・分散モデルを前提に望ましい推定量を提案した研究は，平均ベクトルのスタイン推定量（スタイン推定量は縮小推定量である）を提案したJobson,

[1] Markowitzにより提案された現在の資産運用における標準的な手法。

Korkie and Ratti [4] が唯一と思われる。この提案は残念ながらスタイン推定量の性質[2]からの類推であり，最適化への影響を考慮して求めた結果ではない。資産選択モデルを平均・分散モデルに限らなければ，実は多くの研究が縮小推定量を提案している[3]。しかし，これらの提案には縮小推定量が決定理論の立場から望ましいという以外に理論的根拠はなく，いずれも予想の域を出たものではない。事実，本研究が示したように，条件 $a \geq M \big/ \sqrt{\mathbf{p}'\Sigma^{-1}\mathbf{p}}$ が成立しなければ，縮小推定量が最尤推定量より望ましいとは限らない。

この論文は次のように構成される。次章で二次計画問題への適用を前提とした母数の推定問題を定式化する。ここで漸近的に最尤推定量が非許容的であることと，そのための十分条件が定理として示される。この定理は第３章で証明する。

2 問題の定式化と母数の推定方法

2－1 問題の定式化

本研究が扱う二次計画問題は，正規分布$N_K(\mu, \Sigma)$の母数を係数とした

$$\max\ a\mu'\mathbf{x} - \mathbf{x}'\Sigma\mathbf{x}$$
$$s.t.\ \mathbf{p}'\mathbf{x} = M\ ,\ \mathbf{x} \geq \mathbf{0}$$

である。この二次計画問題を問題Ｉと呼ぶことにする。特に断りのない限り，本研究では問題Ｉの最適解を $\mathbf{x}$ と書くことにする。$\mathbf{x}$ が任意の許容解ではないことに注意する。そして，最適解に対応した目的関数の値は

$$\pi = a\mu'\mathbf{x} - \mathbf{x}'\Sigma\mathbf{x}$$

と書くことにする。

問題Ｉは未知母数 μ, Σ に依存するため，実際には μ, Σ に推定値を代入した二次計画問題が用いられる。この二次計画問題の最適解を $\hat{\mathbf{x}}$ とすれば，この値は問題Ｉの最適解 $\mathbf{x}$ の推定量と考えることができる。最適解 $\hat{\mathbf{x}}$ の下での目的関数は

$$\hat{\pi} = a\mu'\hat{\mathbf{x}} - \hat{\mathbf{x}}'\Sigma\hat{\mathbf{x}} \leq \pi$$

と書くことにする。推定量 $\hat{\mathbf{x}}$ は二次計画問題の最適解と定義されるため，実際に統計的推測の対象となるのは正規分布の母数である。しかし，二次計画問題の目的を考えれば，母数の推定量は真の値からの誤差に基づいて評価すべきではなく，目的関数 $\hat{\pi}$ の誤差に基づいて評価すべきである。そこで，正規分布の母数を推定する評価基準として，本研究では次の平均二乗誤差

$$E(\pi - \hat{\pi})^2 = (\pi - E(\hat{\pi}))^2 + E(\hat{\pi} - E(\hat{\pi}))^2 = \text{バイアス}^2 + \text{分散}$$

を採用する。ただし，通常の推定問題とは異なり，この場合は真の値 π だけでなく推定量に対応する目的関数 $\hat{\pi}$ も未知であることを注意しておく。

---

[2] 多変量正規分布（３次元以上）の平均ベクトルを同時推定する問題で、スタイン推定量は平均二乗誤差の下で最尤推定量を優越するという性質。

[3] Frost and Savarino [2] 、Jorion [5] などが主要な研究である。

本論文冒頭

本研究で基準とする推定方法は最尤推定である。$\mathbf{R}_j$ ；$j=1,\ldots,n$を$N_K(\mu,\Sigma)$ に従う独立な確率変数とすれば，平均ベクトルと分散共分散行列の最尤推定量は

$$\hat{\mu}=\frac{1}{n}\sum_{j=1}^{n}\mathbf{R}_j\ ,\ \hat{\Sigma}=\frac{1}{n}\sum_{j=1}^{n}\left(\mathbf{R}_j-\hat{\mu}\right)'\left(\mathbf{R}_j-\hat{\mu}\right)$$

である。最尤推定量により定義された二次計画問題を**問題**IIと呼ぶことにする。問題IIの最適解は$\hat{\mathbf{x}}$とし，対応する目的関数は$\hat{\pi}$とする。

一方，漸近的に最尤推定量を優越する推定量として，本研究では次の推定量のクラスを取り上げる。そのクラスとは，各成分がゼロに確率収束する統計量 $\hat{\mu}_0$, $\hat{\Sigma}_0$と，1に収束する正の実数列 $\{r_n\},\{s_n\}$ により

$$\tilde{\mu}=r_n\left(\hat{\mu}+\hat{\mu}_0\right),\ \tilde{\Sigma}=s_n\left(\hat{\Sigma}+\hat{\Sigma}_0\right)$$

と定義される推定量のクラスである。明らかに，推定量 $\tilde{\mu},\tilde{\Sigma}$は一致推定量である。これらの推定量により定義される二次計画問題を**問題**IIIとし，その最適解は$\tilde{\mathbf{x}}$，対応する目的関数は$\tilde{\pi}$とする。ただし，$t_n=r_n/s_n$ とすれば，

$$\max\ ar_n\left(\hat{\mu}'+\hat{\mu}_0'\right)\mathbf{x}-\mathbf{x}'s_n\left(\hat{\Sigma}+\hat{\Sigma}_0'\right)\mathbf{x}\Leftrightarrow\max\ at_n\left(\hat{\mu}'+\hat{\mu}_0'\right)\mathbf{x}-\mathbf{x}'\left(\hat{\Sigma}+\hat{\Sigma}_0'\right)\mathbf{x}$$

となるので，実際には $s_n\equiv1$としても一般性を失うことはない。そこで，本研究では実数列 $\{t_n\}$ を前提に議論を進めていくことにする。

ところで，最尤推定量が漸近的に非許容的であることを証明するには，問題Iを標準化した次の二次計画問題を使う必要がある。その二次計画問題とは，$\Sigma=\mathbf{T}'\mathbf{T}$を満たす正則行列$\mathbf{T}$による変数変換$\mathbf{Y}=(\mathbf{T}')^{-1}\mathbf{R}\sim N_K(\nu,\mathbf{I})$に対応した

$$\max\ a\nu'\mathbf{z}-\mathbf{z}'\mathbf{z}$$

$$s.t.\ \mathbf{q}'\mathbf{z}=m\ ,\ \mathbf{T}^{-1}\mathbf{z}\geq\mathbf{0}$$

である。各係数と行列は

$$\nu=\left(\mathbf{T}'\right)^{-1}\mu\ ,\ \mathbf{q}=\frac{\left(\mathbf{T}'\right)^{-1}\mathbf{p}}{\sqrt{\mathbf{p}'\Sigma^{-1}\mathbf{p}}}\ ,\ m=\frac{M}{\sqrt{\mathbf{p}'\Sigma^{-1}\mathbf{p}}}$$

と定義される。この問題の最適解を$\mathbf{z}$とすれば，問題Iの最適解$\mathbf{x}$との間には

$$\mathbf{x}=\mathbf{T}^{-1}\mathbf{z}$$

という関係が成立する。当然，最適化した目的関数の値も，

$$\pi=a\nu'\mathbf{z}-\mathbf{z}'\mathbf{z}=a\mu'\mathbf{x}-\mathbf{x}'\Sigma\mathbf{x}$$

と等しくなる。

本論文冒頭

2-2 母数の推定方法に関する漸近的な結果

二次計画問題への適用を前提とした多変量正規分布の平均ベクトルと分散共分散行列の推定方法について，次の定理が成立する。また，この定理は次章で証明する。

定理　平均二乗誤差の下で問題Ⅱの推定量$\hat{\boldsymbol{\pi}}$を漸近的に優越する問題Ⅲの推定量$\tilde{\boldsymbol{\pi}}$が存在するための十分条件は

$$a \geq M \Big/ \sqrt{\mathbf{p}'\boldsymbol{\Sigma}^{-1}\mathbf{p}}$$

が成立することである。このときの平均ベクトルの推定量は，正数 $\varepsilon_n < 1$，$\varepsilon_n \to 1$ を下限とした縮小因子 $\varepsilon_n < t_n < 1$ により$\tilde{\boldsymbol{\mu}} = t_n\,(\hat{\boldsymbol{\mu}} + \hat{\boldsymbol{\mu}}_0)$と表すことができる。

この定理は二次計画問題が十分条件を満たしていれば，平均ベクトルの望ましい推定量が縮小推定量であることを意味している。既に述べたように，金融工学の分野で最適化問題への適用を前提に推定方法を考察した研究は，望ましい推定量として縮小推定量を提案している。したがって，縮小推定量という予想は本研究が示した定理によって，漸近的な意味ではあるが，理論的に保証されたことになる。そして，この予想が成立するための十分条件も示すことができた。

3 定理の証明

3-1 仮定

前章で提示した定理の証明では，一般性を失うことなく

- 問題Ⅰの最適解について$\mathbf{x} > \mathbf{0}$が成立する。

と仮定する。この仮定により問題Ⅰで非負条件を考える必要がなくなり，以後の議論を簡潔に進めることが可能となる。

一般性を失わないことは次のように示すことができる。Lagrange乗数をλ，スラック変数を$\mathbf{u}' = (u_1 \cdots u_K) \geq \mathbf{0}$とすれば，最適解が満たすKT条件は

$$-2\boldsymbol{\Sigma}\mathbf{x} + \mathbf{u} + \lambda\mathbf{p} + a\boldsymbol{\mu} = \mathbf{0}$$
$$\mathbf{p}'\mathbf{x} - m = 0$$
$$u_k x_k = 0\,;\, \mathit{for\ all}\ k$$

となる。ここで，一般性を失うことなく，

$$x_k = 0 \Leftrightarrow u_k > 0\,;\, k = 1, \ldots, K$$

が成立していると仮定する[4]。いま，$\mathbf{x}' = (\mathbf{x}_1' \ \ \mathbf{x}_2')$と分割したとき，$\mathbf{x}_1 > \mathbf{0}$，$\mathbf{x}_2 = \mathbf{0}$であると

---

[4] $x_k = 0 \Leftarrow u_k > 0$だが、逆が成立するとは限らない。ちょうど、$x_k = u_k = 0$ となるこも起こり得る。

本論文冒頭

しよう。同様に，他のベクトルと行列も

$$\mathbf{u}=\begin{pmatrix}\mathbf{u}_1\\ \mathbf{u}_2\end{pmatrix},\mu=\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix},\Sigma=\begin{pmatrix}\Sigma_{11}&\Sigma_{12}\\ \Sigma_{12}'&\Sigma_{22}\end{pmatrix},\mathbf{p}=\begin{pmatrix}\mathbf{p}_1\\ \mathbf{p}_2\end{pmatrix}$$

と分割する。仮定より$\mathbf{u}_1=\mathbf{0}$，$\mathbf{u}_2>\mathbf{0}$となるから，このときのKT条件

$$-2\Sigma_{11}\mathbf{x}_1+\lambda\mathbf{p}_1+a\mu_1=\mathbf{0}$$
$$\mathbf{p}_1'\mathbf{x}_1-m=0$$
$$-2\Sigma_{12}\mathbf{x}_1+\mathbf{u}_2+\lambda\mathbf{p}_2+a\mu_2=\mathbf{0}$$

から，$\mathbf{x}_1$は$\mu_1,\Sigma_{11},\mathbf{p}_1$によってのみ決められることが分かる。そして，$\mathbf{x}$,$\mathbf{u}$は μ,Σ について連続だから，微小な母数の変化 $\mu+\Delta\mu$, $\Sigma+\Delta\Sigma$に対応した最適解を $\mathbf{x}+\Delta\mathbf{x}$, $\mathbf{u}+\Delta\mathbf{u}$と書くことにすれば，

$$\mathbf{x}_1+\Delta\mathbf{x}_1>\mathbf{0}\,,\,\mathbf{u}_2+\Delta\mathbf{u}_2>\mathbf{0}\Rightarrow\mathbf{x}_2+\Delta\mathbf{x}_2=\mathbf{0}\,,\,\mathbf{u}_1+\Delta\mathbf{u}_1=\mathbf{0}$$

がやはり成立する。このように，最適解の μ,Σ による摂動を調べるには，$x_k>0$ となる変数についてのみ調べれば十分である

定理の証明は問題IIとIIIの最適解 $\hat{\mathbf{x}},\tilde{\mathbf{x}}$ の漸近分布に基づいて進められる。二つの問題に対応した正規母数の推定量はいずれも一致推定量だから，最適解の漸近分布を調べるということは，最適解の真の母数による摂動を調べることに外ならない。以上が問題Iの最適解$\mathbf{x}$を正と仮定して一般性を失わない理由である。

3-2 最適解の漸近分布

準備として以下のベクトルと行列を定義しておく。まず，分散共分散行列の非対角成分は重複しているので，ここでは $K(K+1)/2$個の非重複部分を並べたベクトルを

$$\boldsymbol{\sigma}'=\left(\sigma_{11}\cdots\sigma_{1K}\ \sigma_{22}\cdots\sigma_{2K}\cdots\sigma_{K-1,K-1}\ \sigma_{K-1,K}\ \sigma_{KK}\right)$$

と書くことにする。次に，n次元ベクトル$\mathbf{x}$に対して$n\times n(n+1)/2$型行列 $\mathbf{W}_{\mathbf{x}}$を

$$\mathbf{W}_{\mathbf{x}}=\begin{pmatrix}x_1\cdots x_n&&&0\\ &x_2\cdots x_n&&\\ &&\ddots&\\ &0&&x_n\end{pmatrix}$$

と定義する。行列$\mathbf{W}_{\mathbf{x}}$は明らかに線形性を持つ。

さて，定理の証明は $t_n\equiv t\cong1$ と固定して議論を進めていく。そして，$t<1$ ならば漸近分布の下で推定量 $\tilde{\mu},\tilde{\boldsymbol{\sigma}}$ が最尤推定量を優越することと，このようなtの下限が標本数 n に依存することは証明の最後に示される。

まず，問題IIIで $\hat{\mu}_0\equiv0,\hat{\Sigma}_0\equiv0$，すなわち，$\tilde{\mu}=t\hat{\mu},\tilde{\Sigma}=\hat{\Sigma}$とした場合の最適解$\tilde{\mathbf{x}}$の漸近分布から考えていく。問題Iで係数 $ta,t\cong1$ をと変化させたときの最適解を$\mathbf{x}_t$と書くことにすれば，仮定から $t=1$ の近傍で非負条件は不要だから，この値はLagrange乗数法を用いて求めることが

本論文冒頭

できる。そこで，

$$\mathbf{H}=\Sigma^{-1}\left(\mathbf{I}-s\mathbf{pp}'\Sigma^{-1}\right),\ s=1/\mathbf{p}'\Sigma^{-1}\mathbf{p}$$

とすれば，最適解は

$$\mathbf{x}_t=\frac{at}{2}\mathbf{H}\mu+sM\Sigma^{-1}\mathbf{p}$$

となる。 $t=1$のときの最適解$\mathbf{x}_t$は問題 I の最適解$\mathbf{x}$となるから，

$$\mathbf{x}_t=\mathbf{x}+\frac{a}{2}(t-1)\mathbf{H}\mu$$

が成立する。最適解 $\mathbf{x}_t$の母数 μ, Σ による偏導関数を行列 $\mathbf{G}_t$で表せば，行列 $\mathbf{G}_t$は

$$\mathbf{G}_t=\begin{pmatrix}\dfrac{\partial\mathbf{x}_t}{\partial\mu'} & \dfrac{\partial\mathbf{x}_t}{\partial\sigma'}\end{pmatrix}$$
$$=\begin{pmatrix}\dfrac{at}{2}\mathbf{H} & -\mathbf{HW}_{\mathbf{x}_t}\end{pmatrix}=\begin{pmatrix}\dfrac{at}{2}\mathbf{H} & -\mathbf{HW}_{\mathbf{x}}-\dfrac{a}{2}(t-1)\mathbf{HW}_{\mathbf{H}\mu}\end{pmatrix}$$

となる。この導出は付録 1 に載せてある。行列 $\mathbf{G}_t$は，次の行列

$$\mathbf{G}=\mathbf{G}_t\big|_{t=1}=\begin{pmatrix}\dfrac{a}{2}\mathbf{H} & -\mathbf{HW}_{\mathbf{x}}\end{pmatrix},\ \mathbf{B}=\begin{pmatrix}\dfrac{a}{2}\mathbf{H} & -\dfrac{a}{2}\mathbf{HW}_{\mathbf{H}\mu}\end{pmatrix}$$

により，

$$\mathbf{G}_t=\mathbf{G}+(t-1)\mathbf{B}$$

と表すことができる。

最尤推定量 $\tilde{\mu}, \tilde{\sigma}$ の漸近分布を

$$\sqrt{n}\left(\begin{pmatrix}\hat{\mu}\\ \hat{\sigma}\end{pmatrix}-\begin{pmatrix}\mu\\ \sigma\end{pmatrix}\right)\xrightarrow{L}N_{K+K(K+1)/2}(\mathbf{0},\mathbf{D})$$

と書くことにすれば，問題 III の最適解 $\tilde{\mathbf{x}}$ の漸近分布は

$$\sqrt{n}\left(\tilde{\mathbf{x}}-\left\{\mathbf{x}+\frac{a}{2}(t-1)\mathbf{H}\mu\right\}\right)\xrightarrow{L}N_K\left(\mathbf{0},(\mathbf{G}+(t-1)\mathbf{B})\mathbf{D}(\mathbf{G}'+(t-1)\mathbf{B}')\right)$$

となる[5]。ここで，$t=1$とすれば問題 II の最適解 $\hat{\mathbf{x}}$ の漸近分布となる。すなわち，

$$\sqrt{n}(\hat{\mathbf{x}}-\mathbf{x})\xrightarrow{L}N_K(\mathbf{0},\mathbf{GDG}')$$

である。いずれの漸近分布も分散共分散行列の階数は$K-1$だから，これらの多変量正規分布は密度関数が存在しない。$P(\mathbf{p}'\hat{\mathbf{x}}=M)=P(\mathbf{p}'\tilde{\mathbf{x}}=M)=1$という性質から，これは当然の結果である。

---

[5] Rao ［6］, p388, (iii) を用いた。

本論文冒頭

最後に推定量$\tilde{\mu} = t(\hat{\mu} + \hat{\mu}_0)$, $\tilde{\Sigma} = \hat{\Sigma} + \hat{\Sigma}_0$についての議論に戻ると，結論から言えば，対応する最適解 $\tilde{\mathbf{x}}$ と推定量$\tilde{\mu} = t\hat{\mu}, \tilde{\Sigma} = \hat{\Sigma}$に対応する最適解は同じ漸近分布を持つ。これは後者の漸近分布を求めた議論と，次の二つの性質から明らかである。まず，

$$(\hat{\mu} + \hat{\mu}_0) - \hat{\mu} = \hat{\mu}_0 \xrightarrow{P} \mathbf{0}, \ (\hat{\Sigma} + \hat{\Sigma}_0) - \hat{\Sigma} = \hat{\Sigma}_0 \xrightarrow{P} \mathbf{0}$$

が成立するから，$\tilde{\mu} + \hat{\mu}_0, \tilde{\Sigma} + \hat{\Sigma}_0$は最尤推定量と同じ漸近分布を持つことである。二点目は，

$\tilde{\mu} \xrightarrow{P} t\mu, \ \tilde{\Sigma} \xrightarrow{P} \Sigma$ だから，既に定義した行列$\mathbf{G}_t$を，そのまま適用できるということである。

3－3　定理の証明

平均二乗誤差はバイアスの二乗と分散の和に分解される。ここでは

$$\hat{\pi}, \tilde{\pi} \le \pi \Rightarrow E(\hat{\pi}), E(\tilde{\pi}) \le \pi$$

が成立するから，期待値の大小とバイアスの二乗の大小は一致する。したがって，各推定量の漸近分布の下で

$$\varepsilon_n < t < 1 \Rightarrow E(\tilde{\pi}) > E(\hat{\pi}), \ V(\tilde{\pi}) < V(\hat{\pi})$$

を示せば，$E(\tilde{\pi} - \pi)^2 < E(\hat{\pi} - \pi)^2$となり，定理を証明したことになる。

まず，$E(\tilde{\pi}) > E(\hat{\pi})$を証明する。問題Ⅲにおける目的関数 $\tilde{\pi} = a\mu'\tilde{\mathbf{x}} - \tilde{\mathbf{x}}'\Sigma\tilde{\mathbf{x}}$の漸近分布に関する期待値は

$$E(\tilde{\pi}) = a\mu'\left(\mathbf{x} + \frac{a}{2}(t-1)\mathbf{H}\mu\right) - \left(\mathbf{x}' + \frac{a}{2}(t-1)\mu'\mathbf{H}\right)\Sigma\left(\mathbf{x} + \frac{a}{2}(t-1)\mathbf{H}\mu\right)$$
$$- \frac{1}{n}\mathrm{Tr}\,\Sigma(\mathbf{G} + (t-1)\mathbf{B})\mathbf{D}(\mathbf{G}' + (t-1)\mathbf{B}')$$

である。ここで，

$$\mu'\mathbf{H}\Sigma\mathbf{H}\mu = \mu'\mathbf{H}\mu, \ \mathbf{x}'\Sigma\mathbf{H}\mu = \frac{a}{2}\mu'\mathbf{H}\mu$$

に注意して，$t-1$ の項だけ求めると，この期待値は

$$E(\tilde{\pi}) = -\frac{2}{n}(t-1)\,\mathrm{Tr}\,\Sigma\mathbf{GDB}' + (*)$$

となる。付録２で示すように，$\mathrm{Tr}\Sigma\mathbf{GDB}' > 0$ だから，

$$\left.\frac{\partial}{\partial t}E(\tilde{\pi})\right|_{t=1} = -\frac{2}{n}\mathrm{Tr}\,\Sigma\mathbf{GDB}' < 0$$

本論文冒頭

が成立する。これは期待値$E(\tilde{\pi})$が $t=1$ の近傍で t の減少関数ということである。この結果から，標本数nに応じて正数$\zeta_n<1$, $\zeta_n \to 1$が存在し，$\zeta_n<t<1$ならば漸近的に

$$E(\tilde{\pi}) > E(\tilde{\pi})\big|_{t=1} = E(\hat{\pi})$$

であることが分かる。

次に，$V(\tilde{\pi})<V(\hat{\pi})$を証明する。ここでは二次形式の分散に関する

$$\mathbf{X} \sim N(\mu, \Sigma),\ \mathbf{A}=\mathbf{A}'$$
$$\Rightarrow V\{(\mathbf{X}'-\mathbf{a}')\mathbf{A}(\mathbf{X}-\mathbf{a})\} = 2\,\mathrm{Tr}\,(\mathbf{A}\Sigma)^2 + 4(\mu-\mathbf{a})'\mathbf{A}\Sigma\mathbf{A}(\mu-\mathbf{a})$$

という結果を用いる。$\Sigma=\mathbf{T}'\mathbf{T}$に注意すると，問題Ⅲにおける目的関数$\tilde{\pi}$は

$$\begin{aligned}\tilde{\pi} &= a\mu'\tilde{\mathbf{x}} - \tilde{\mathbf{x}}'\Sigma\tilde{\mathbf{x}} \\ &= -\left(\tilde{\mathbf{x}}'\mathbf{T}' - \frac{1}{2}a\mu'\mathbf{T}^{-1}\right)\left(\mathbf{T}\tilde{\mathbf{x}} - \frac{1}{2}a(\mathbf{T}')^{-1}\mu\right) + \frac{1}{4}a^2\mu'\Sigma^{-1}\mu\end{aligned}$$

と書くことができる。漸近分布における$\mathbf{T}\tilde{\mathbf{x}}$の期待値と分散は

$$\begin{aligned}E(\mathbf{T}\tilde{\mathbf{x}}) &= \frac{at}{2}\mathbf{T}\mathbf{H}\mu + sM\mathbf{T}\Sigma^{-1}\mathbf{p} \\ V(\mathbf{T}\tilde{\mathbf{x}}) &= \mathbf{T}V(\tilde{\mathbf{x}})\mathbf{T}' \\ &= \frac{1}{n}\mathbf{T}(\mathbf{G}+(t-1)\mathbf{B})\mathbf{D}(\mathbf{G}'+(t-1)\mathbf{B}')\mathbf{T}'\end{aligned}$$

である。したがって，$\mathbf{p}'\mathbf{G}=\mathbf{p}'\mathbf{B}=\mathbf{0}$, $\mathbf{H}\Sigma\mathbf{G}=\mathbf{G}$, $\mathbf{H}\Sigma\mathbf{B}=\mathbf{B}$ に注意すれば，

$$\begin{aligned}V(\tilde{\pi}) &= 2\,\mathrm{Tr}\left\{\frac{1}{n}\mathbf{T}(\mathbf{G}+(t-1)\mathbf{B})\mathbf{D}(\mathbf{G}'+(t-1)\mathbf{B}')\mathbf{T}'\right\}^2 \\ &\quad + 4\left(\frac{a}{2}\mu'(t\mathbf{H}\mathbf{T}'-\mathbf{T}^{-1}) + sM\mathbf{p}'\Sigma^{-1}\mathbf{T}'\right) \\ &\quad \times \left\{\frac{1}{n}\mathbf{T}(\mathbf{G}+(t-1)\mathbf{B})\mathbf{D}(\mathbf{G}'+(t-1)\mathbf{B}')\mathbf{T}'\right\}\left(\frac{a}{2}(t\mathbf{T}\mathbf{H}-(\mathbf{T}')^{-1})\mu + sM\mathbf{T}\Sigma^{-1}\mathbf{p}\right) \\ &= \frac{2}{n^2}\mathrm{Tr}\,\Sigma\{(\mathbf{G}+(t-1)\mathbf{B})\mathbf{D}(\mathbf{G}'+(t-1)\mathbf{B}')\}^2 \\ &\quad + \frac{a^2}{4n}(t-1)^2\mu'(\mathbf{G}+(t-1)\mathbf{B})\mathbf{D}(\mathbf{G}'+(t-1)\mathbf{B}')\mu\end{aligned}$$

であることが分かる。この値を$t-1$ の項だけ求めると，

本論文冒頭

$$
\begin{aligned}
V(\tilde{\pi}) &= \frac{2}{n^2}\mathrm{Tr}\,\Sigma\{(\mathbf{G}+(t-1)\mathbf{B})\mathbf{D}(\mathbf{G}'+(t-1)\mathbf{B}')\}^2+(*)\\
&= \frac{2}{n^2}\mathrm{Tr}\,\Sigma\{\mathbf{GDG}'+(t-1)(\mathbf{GDB}'+\mathbf{BDG}')+(t-1)^2\mathbf{BDB}'\}^2+(*)\\
&= \frac{2}{n^2}\mathrm{Tr}\,\Sigma\{(t-1)(\mathbf{GDG}'(\mathbf{GDB}'+\mathbf{BDG}')+(\mathbf{GDB}'+\mathbf{BDG}')\mathbf{GDG}')+(*)\}+(*)\\
&= \frac{8}{n^2}(t-1)\mathrm{Tr}\,\Sigma\mathbf{GDG}'\mathbf{GDB}'+(*)
\end{aligned}
$$

となる。付録３で示すように，$\mathrm{Tr}\Sigma\mathbf{GDG}'\mathbf{GDB}' > 0$ だから，次の不等式が成立する。

$$
\left.\frac{\partial}{\partial t}V(\tilde{\pi})\right|_{t=1} = \frac{8}{n^2}\mathrm{Tr}\,\Sigma\mathbf{GDG}'\mathbf{GDB}' > 0
$$

この結果は，正数$\eta_n < 1$, $\eta_n \to 1$が存在し，$\eta_n < t < 1$ならば漸近的に

$$
V(\tilde{\pi}) < V(\tilde{\pi})\big|_{t=1} = V(\hat{\pi})
$$

で　ることを意味している。

そこで，正数 ε_n を$\varepsilon_n = \max\{\zeta_n, \eta_n\} < 1$と定義すれば，$\varepsilon_n \to 1$ であり，漸近的に

$$
\varepsilon_n < t < 1 \Rightarrow E(\tilde{\pi}) > E(\hat{\pi}),\ V(\tilde{\pi}) < V(\hat{\pi})
$$

となる。したがって，漸近的に

$$
E(\tilde{\pi}-\pi)^2 < E(\hat{\pi}-\pi)^2
$$

が成立する。実数列 $\{t_n\}$ は条件 $\varepsilon_n < t_n < 1; n = 1, 2, \cdots$を満たすように選べばよい。以上により定理は証明された。

本論文冒頭

4 まとめ

本研究では，二次計画問題の係数として多変量正規分布の平均ベクトルと分散共分散行列を推定する問題について考察した。そして，ある条件が成立しているならば，慣例的に用いられている最尤推定量が漸近的に非許容的であることを証明した。平均ベクトルの縮小推定量が最尤推定量を優越することも示すことができた。これは既存研究の予想と一致している。したがって，本研究は既存研究の予想を漸近的な意味において正当化し，その十分条件を明らかにしたという意味で詳細に記述したことになる。

もちろん，どのような推定値を用いるべきかという問いに対して，この定理は不十分と言わざるを得ない。最尤推定量が漸近的に非許容的であるという主張だけで，縮小因子の下限 ε_n が具体的に示されていないからである。この値を評価するには係数 t の四次式となる分散の評価式 $V(\tilde{\pi})<V(\hat{\pi})$ を代数的に調べる必要があり，これが未解決の問題として残されている[6]。さらに，標本数に制限のある現実の問題では漸近的性質にあまり意味がないことも指摘する必要があろう。

今後の課題は，上で述べたように縮小因子の下限を正確に評価することと，縮小推定量が望ましいという主張を小標本下で検証することである。また，定理では二次計画問題の係数として正規分布の母数を仮定しているが，この仮定に関する頑健性も興味深い課題と言うことができる。

〈付録1〉

ここでは行列 $\mathbf{G}_t$ の導出を説明する。問題 I で係数を $a \to ta,\ t \cong 1$ と変化させた二次計画問題の Lagrange関数 L と，その偏導関数は，Lagrange乗数を λ とすれば，

$$L = at\mu'\mathbf{x}_t - \mathbf{x}_t'\Sigma\mathbf{x}_t + \lambda(\mathbf{p}'\mathbf{x}_t - M)$$

$$\frac{\partial L}{\partial \mathbf{x}_t} = at\mu - 2\Sigma\mathbf{x}_t + \lambda\mathbf{p}$$

$$\frac{\partial L}{\partial \lambda} = \mathbf{p}'\mathbf{x}_t - M$$

である。$\partial L/\partial \mathbf{x}_t$ を $L_1 \cdots L_K$，$\partial L/\partial \lambda$ を L_{K+1} とすれば，各 L_K の偏導関数行列は

$$\frac{\partial \mathbf{L}}{\partial(\mathbf{x}_t,\lambda)} = \begin{pmatrix} \dfrac{\partial L_{1:K}}{\partial \mathbf{x}_t'} & \dfrac{\partial L_{1:K}}{\partial \lambda} \\ \dfrac{\partial L_{K+1}}{\partial \mathbf{x}_t'} & \dfrac{\partial L_{K+1}}{\partial \lambda} \end{pmatrix} = \begin{pmatrix} -2\Sigma & \mathbf{p} \\ \mathbf{p}' & 0 \end{pmatrix},\ \frac{\partial \mathbf{L}}{\partial(\mu,\sigma)} = \begin{pmatrix} \dfrac{\partial L_{1:K}}{\partial \mu'} & \dfrac{\partial L_{1:K}}{\partial \sigma'} \\ \dfrac{\partial L_{K+1}}{\partial \mu'} & \dfrac{\partial L_{K+1}}{\partial \sigma'} \end{pmatrix} = \begin{pmatrix} at\mathbf{I} & \mathbf{W}_{\mathbf{x}_t} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

[6] 期待値に関する下限が $\zeta_n = 1 - 8\,\mathrm{Tr}\,\mathbf{\Sigma GDB}'/(na^2\mu'\mathbf{H}\mu + 4\,\mathrm{Tr}\,\mathbf{\Sigma BDB}')$ であることは、評価式 $E(\tilde{\pi})>E(\hat{\pi})$ から容易に求めることができる。

本論文冒頭

となる。この結果に陰関数定理

$$\frac{\partial(\mathbf{x}_t,\lambda)}{\partial(\mu,\sigma)}=\begin{pmatrix}\dfrac{\partial\mathbf{x}_t}{\partial\mu'} & \dfrac{\partial\mathbf{x}_t}{\partial\sigma'}\\ \dfrac{\partial\lambda}{\partial\mu'} & \dfrac{\partial\lambda}{\partial\sigma'}\end{pmatrix}=-\left(\frac{\partial\mathbf{L}}{\partial(\mathbf{x}_t,\lambda)}\right)^{-1}\frac{\partial\mathbf{L}}{\partial(\mu,\sigma)}$$

を適用すれば，最適解$\mathbf{x}_t$の母数μ,Σによる偏導関数行列$\mathbf{G}_t$を求めることができる。右辺の逆行列は各小行列により

$$\left(\frac{\partial\mathbf{L}}{\partial(\mathbf{x}_t,\lambda)}\right)^{-1}=\begin{pmatrix}-2\Sigma & \mathbf{p}\\ \mathbf{p}' & 0\end{pmatrix}^{-1}=\begin{pmatrix}-\mathbf{H}/2 & s\Sigma^{-1}\mathbf{p}\\ s\mathbf{p}'\Sigma^{-1} & 2s\end{pmatrix}$$

と表される。ただし，$\mathbf{H}=\Sigma^{-1}\left(\mathbf{I}-s\mathbf{p}\mathbf{p}'\Sigma^{-1}\right)$, $s=1/\mathbf{p}'\Sigma^{-1}\mathbf{p}$ である。

〈付録2〉

不等式$\mathrm{Tr}\Sigma\mathbf{GDB}'>0$を証明する。ここでは標準化された問題Ｉを用いるので，各行列や係数は

$$\mathbf{x}\to\mathbf{z}\,,\ \mu\to\nu\,,\ \Sigma\to\mathbf{I}\,,\ \mathbf{p}\to\mathbf{q}\,,\ M\to m$$

となる。最尤推定量$\hat{\mu},\hat{\sigma}$の漸近分散は，正規分布が$N_K(\nu,\mathbf{I})$の場合は

$$\frac{1}{n}\mathbf{D}=\frac{1}{n}\begin{pmatrix}\mathbf{I}_K & \mathbf{0}\\ \mathbf{0} & \Lambda\end{pmatrix},\ \Lambda=\begin{pmatrix}\Lambda_{(1)} & & \mathbf{0}\\ & \ddots & \\ \mathbf{0} & & \Lambda_{(K)}\end{pmatrix},\ \Lambda_{(k)}=4\begin{pmatrix}1/3 & \mathbf{0}\\ \mathbf{0} & \mathbf{I}_{k-1}\end{pmatrix}$$

である。行列$\mathbf{H}=\mathbf{I}-\mathbf{q}\mathbf{q}'$が階数 $K-1$ の直行射影行列であることに注意すると，

$$\mathrm{Tr}\,\mathbf{GDB}'=\mathrm{Tr}\begin{pmatrix}\dfrac{a}{2}\mathbf{H} & -\mathbf{HW_z}\end{pmatrix}\begin{pmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \Lambda\end{pmatrix}\begin{pmatrix}\dfrac{a}{2}\mathbf{H}\\ -\dfrac{a}{2}\mathbf{W}'_{\mathbf{H}\nu}\mathbf{H}\end{pmatrix}$$

$$=\frac{a^2}{4}(K-1)+\frac{a}{2}\mathrm{Tr}\,\mathbf{W_z}\Lambda\mathbf{W}'_{\mathbf{H}\nu}\mathbf{H}$$

本論文冒頭

なる。対角行列Λのサイズを

$$\Lambda \to \tilde{\Lambda} = \begin{pmatrix} \tilde{\Lambda}_{(1)} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \tilde{\Lambda}_{(K)} \end{pmatrix}, \tilde{\Lambda}_{(k)} = \begin{pmatrix} \mathbf{0}_{K-k} & \mathbf{0} \\ \mathbf{0} & \Lambda_{(k)} \end{pmatrix}$$

と拡大すれば，第二項は次のように書くことができる。

$$\begin{aligned} &\mathrm{Tr}\ \mathbf{W}_{\mathbf{z}}\Lambda\mathbf{W}'_{\mathbf{H}\nu}\mathbf{H} \\ &= \mathrm{Tr}\begin{pmatrix} \mathbf{z}' & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{z}' \end{pmatrix}\begin{pmatrix} \tilde{\Lambda}_{(1)} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \tilde{\Lambda}_{(K)} \end{pmatrix}\begin{pmatrix} \mathbf{H}\nu & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{H}\nu \end{pmatrix}\begin{pmatrix} 1-q_1^2 & & -q_1q_K \\ & \ddots & \\ -q_Kq_1 & & 1-q_K^2 \end{pmatrix} \\ &= \sum_{k=1}^{K}\left(1-q_k^2\right)\mathbf{z}'\tilde{\Lambda}_{(k)}\mathbf{H}\nu \\ &= \left(\frac{a}{2}\nu'\mathbf{H}+m\mathbf{q}'\right)\left\{\sum_{k=1}^{K}\left(1-q_k^2\right)\tilde{\Lambda}_{(k)}\right\}\mathbf{H}\nu \end{aligned}$$

最後の二次形式にある対角行列を

$$\begin{aligned} \Theta &= \mathrm{diag}\left(\left(1-q_1^2\right)-\frac{2}{3}\left(1-q_1^2\right) \quad \cdots \quad \sum_{k=1}^{K}\left(1-q_k^2\right)-\frac{2}{3}\left(1-q_K^2\right)\right) \\ &= \mathrm{diag}\left(\theta_1 \quad \cdots \quad \theta_K\right) \end{aligned}$$

とおく。$\sum q_k^2 = 1$ に注意すると，$\theta_1 \le \cdots \le \theta_K$であり，

$$k-\frac{5}{3} < \theta_k < k-\frac{2}{3}$$

であることが分かる。ただし，ここでは$0 \le \theta_1 < 1/3$となるように各変数は並んでいるものとする（一般には$0 \le \theta_1 < 1/3$である）。

さて，$\nu \in \mathbf{R}^K$であるが，行列$\mathbf{H}$は$\mathbf{R}^K \to V(\mathbf{q})^{\perp}$への射影行列だから，$\xi = \mathbf{H}\nu$とすれば，このベクトルは$\mathbf{q}'\xi = 0$と特徴付けることができる。そこで，この制約の下で，

$$\mathrm{Tr}\ \mathbf{W}_{\mathbf{z}}\Lambda\mathbf{W}'_{\mathbf{H}\nu}\mathbf{H} = \left(\frac{a}{2}\nu'\mathbf{H}+m\mathbf{q}'\right)\Theta\mathbf{H}\nu = m\mathbf{q}'\Theta\xi+\frac{a}{2}\xi'\Theta\xi$$

本論文冒頭

の最小値を求めると，その値は

$$\min_{\mathbf{q}'\xi=0} m\mathbf{q}'\Theta\xi+\frac{a}{2}\xi'\Theta\xi=-\frac{m^2}{4}\left(\mathbf{q}'\Theta\mathbf{q}-\frac{1}{\mathbf{q}'\Theta^{-1}\mathbf{q}}\right)$$

となる。この結果を$\mathrm{Tr}\mathbf{GDB}'$に代入すると次のようになる。

$$\begin{aligned}\mathrm{Tr}\,\mathbf{GDB}'&=\frac{a^2}{4}(K-1)+\frac{a}{2}\left(m\mathbf{q}'\Theta\xi+\frac{a}{2}\xi'\Theta\xi\right)\\&\geq\frac{a^2}{4}(K-1)-\frac{m^2}{4}\left(\mathbf{q}'\Theta\mathbf{q}-\frac{1}{\mathbf{q}'\Theta^{-1}\mathbf{q}}\right)\end{aligned}$$

行列Θの最大固有値は θ_Kであり，最小固有値は$\theta_1>0$だから，

$$\mathbf{q}'\Theta\mathbf{q}-\frac{1}{\mathbf{q}'\Theta^{-1}\mathbf{q}}<\theta_K-\frac{1}{1/\theta_1}=\theta_K-\theta_1<K-1$$

となる。固有値θ_1,θ_Kはベクトル$\mathbf{q}$の関数であるが，最後の不等式は任意の$\mathbf{q}$に対して成立することを注意しておく。最後に，定理の仮定$a\geq m$を使えば，

$$\mathrm{Tr}\,\mathbf{GDB}'>\frac{a^2}{4}(K-1)-\frac{m^2}{4}(K-1)=\frac{1}{4}(a^2-m^2)(K-1)>0$$

であることが分かる。

〈付録3〉

ここでは不等式$\mathrm{Tr}\Sigma\mathbf{GDG}'\mathbf{GDB}'>0$を証明する。証明では付録2と同様に，標準化された問題Ⅰを用いる。不等式の各行列を小行列により表すと

$$\begin{aligned}&\mathrm{Tr}\,\mathbf{GDG}'\mathbf{GDB}'\\&=\mathrm{Tr}\left(\frac{a^4}{16}\mathbf{H}+\frac{a^2}{4}\mathbf{HW_z\Lambda W_z'H}+\frac{a^3}{8}\mathbf{HW_z\Lambda W_{Hv}'H}+\frac{a}{2}\mathbf{HW_z\Lambda W_z'HW_z\Lambda W_{Hv}'H}\right)\\&=\frac{a^2}{4}\left(\frac{a^2}{4}\mathrm{Tr}\,\mathbf{H}+\frac{a}{2}\mathrm{Tr}\,\mathbf{HW_z\Lambda W_{Hv}'H}\right)\\&\quad+\frac{a^2}{4}\mathrm{Tr}\,\mathbf{HW_z\Lambda W_{Hv}'H}+\frac{a}{2}\mathrm{Tr}\,\mathbf{W_z\Lambda W_{Hv}'(HW_z\Lambda W_{Hv}'H)}\end{aligned}$$

本論文冒頭

となる。付録２で示したように第一項は正だから，残りの部分が正であることを示せば証明は完了する。まず，

$$\mathbf{W}_{\mathbf{z}}\Lambda\mathbf{W}_{\mathbf{z}}' = \begin{pmatrix} \mathbf{z}'\widetilde{\Lambda}_{(1)}\mathbf{z} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{z}'\widetilde{\Lambda}_{(K)}\mathbf{z} \end{pmatrix}$$

に注意すれば，$\mathbf{HW_z\Lambda W_z'H}$の対角成分が正であることは明らかである。そこで，この対角成分を $\psi_k ; k=1,\ldots,K$ と書くことにする。第三項（のトレース部分）は

$$\mathrm{Tr}\ \mathbf{W}_{\mathbf{z}}\Lambda\mathbf{W}_{\mathbf{H}\nu}'(\mathbf{HW}_{\mathbf{z}}\Lambda\mathbf{W}_{\mathbf{H}\nu}'\mathbf{H}) = \sum_{k=1}^{K}\psi_k\mathbf{z}'\widetilde{\Lambda}_{(k)}\mathbf{H}\nu = \mathbf{z}'\left\{\sum_{k=1}^{K}\psi_k\widetilde{\Lambda}_{(k)}\right\}\mathbf{H}\nu$$

となるから，付録２と同様にして評価することができる。そこで，

$$\Omega = \mathrm{diag}\left(\psi_1 - \frac{2}{3}\psi_1 \quad \cdots \quad \sum_{k=1}^{K}\psi_k - \frac{2}{3}\psi_K\right) = \mathrm{diag}(\omega_1 \quad \cdots \quad \omega_K)$$

とおけば，$0<\omega_1<\cdots<\omega_K$ だから，

$$\mathbf{z}'\left\{\sum_{k=1}^{K}\psi_k\widetilde{\Lambda}_{(k)}\right\}\mathbf{H}\nu = \mathbf{z}'\Omega\mathbf{H}\nu \geq -\frac{m^2}{2a}\left(\mathbf{q}'\Omega\mathbf{q} - \frac{1}{\mathbf{q}'\Omega^{-1}\mathbf{q}}\right) \geq -\frac{m^2}{2a}(\omega_K - \omega_1)$$

であることが分かる。以上から，

$$\frac{a^2}{4}\mathrm{Tr}\ \mathbf{HW}_{\mathbf{z}}\Lambda\mathbf{W}_{\mathbf{H}\nu}'\mathbf{H} + \frac{a}{2}\mathrm{Tr}\ \mathbf{W}_{\mathbf{z}}\Lambda\mathbf{W}_{\mathbf{H}\nu}'(\mathbf{HW}_{\mathbf{z}}\Lambda\mathbf{W}_{\mathbf{H}\nu}'\mathbf{H})$$
$$\geq \frac{a^2}{4}\sum_{k=1}^{K}\psi_k - \frac{m^2}{4}(\omega_K - \omega_1) = \frac{a^2}{4}\left(\omega_K + \frac{2}{3}\psi_K\right) - \frac{m^2}{4}(\omega_K - \omega_1) > 0$$

が成立する。ゆえに，$\mathrm{Tr}\Sigma\mathbf{GDG'GDB'} > 0$ であることが証明された。

参考文献

[1] Britten-Jones, M.(1999), The Sampling Error in Estimates of Mean-Variance Efficient Portfolio Weights, *Journal of Finance*, **54**, 655-671.

[2] Frost, P. A. and Savarino, J. E.(1986). An Empirical Bayes Approach to Efficient Portfolio Selection, *Journal of Financial and Quantitative Analysis.* **21**, 293-305.

[3] Jobson, J. D. and Korkie, B.(1980). Estimation for Markowitz Efficient Portfolios, *Journal of*

本論文冒頭

The American Statistical Association. **75**, 544-554.

[4] Jobson, J. D., Korkie, B. and Ratti, V.(1979). Improved Estimation for Markowitz Portfolios using James-Stein Type Estimators, *Journal of The American Statistical Association.* Business and Economics Statistics Section, **41**, 279-284.

[5] Jorion, P.(1986). Bayes-Stein Estimation for Portfolio Analysis, *Journal of Financial and Quantitative Analysis.* **21**, 279-292.

[6] Rao, C.R.(1973). *Linear Statistical Inference and Its Applications*, 2nd ed, John Wiley.

本論文冒頭