The Burmeister-McElroy Sliced Normal Theorem

Conditional forecasting with probabilistic scenarios

by

Edwin Burmeister
and
Marjorie B. McElroy

Department of Economics
Box 90097
Duke University
Durham, NC 27708-0097

September 24, 1997
Revised April 19, 1999
Revised October 3, 2002

1. The problem

A conditional forecast produces the mean and variance of return, conditional on the forecasted variables that impact return via a linear factor model. These conditional forecasts are easy to compute because the historical factor realizations are approximately multivariate normal.

However, this is not the way in which many people think about forecasting. Rather, they want to say, "I think there is a 50% chance bond yields will rise by 50 b.p., a 30% chance they will stay the same, and a 20% chance they will fall by 10 b.p."

We will call such forecasts ***probabilistic scenarios*** to distinguish them from a pure conditional forecast. We need to compute the joint multivariate distribution for all the factors impacting impact return when the user forecasts some of them with a probabilistic scenario. To do so we need a new result, the ***Sliced Normal Theorem*** proved below.

2. Technical background information and notation

In this section we state well-known results and establish notation. One reference for these results is ***Introduction to the Theory of Statistics*** by Alexander M. Mood and Franklin A. Graybill (New York: McGraw-Hill Book Company, Inc., Second Edition, 1963), Chapter 9, pages 198-219.

The random vector $\mathbf{Y}$ *is distributed as the p-variate normal if the joint density of* $y_1, y_2, \ldots, y_p$ *is*

$$h(\mathbf{Y}) = h(y_1, y_2, \ldots, y_p) = \frac{|R|^{\frac{1}{2}}}{(2\pi)^{\frac{p}{2}}} e^{-\frac{1}{2}(Y-\mu)' R(Y-\mu)} \quad \text{for} \quad -\infty < y_i < +\infty \;\; \text{and} \;\; i = 1, 2, \ldots, p$$

and where

(a) R is a positive definite matrix whose elements r_{ij} *are constants, and*

(b) μ *is a* $p \times 1$ *vector whose elements* μ_i *are constants.*

The univariate case with p=1 is obtained by setting $r_{11} = \frac{1}{\sigma^2}$. The quantity $Q = (Y - \mu)' R(Y - \mu)$ is called the quadratic form of the p-variate normal. It is a theorem that

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(Y-\mu)' R(Y-\mu)} dy_1 \cdots dy_p = (2\pi)^{p/2} |R|^{-1/2}$$

.

and does not depend on the vector μ.

The $p \times p$ covariance matrix of the y 's is

$$V = \begin{bmatrix} \sigma_{\mathbf{11}} & \cdots & \sigma_{\mathbf{1p}} \\ \vdots & & \vdots \\ \sigma_{\mathbf{p1}} & \cdots & \sigma_{\mathbf{pp}} \end{bmatrix}.$$

It is also a theorem (Theorem 9.9, page 211) that $V = R^{-1}$.

We now define the following partitions:

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \quad \mu = \begin{pmatrix} U_1 \\ U_2 \end{pmatrix} \quad R = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \quad \mathrm{V} = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}$$

where

$$Y_1 = \begin{pmatrix} y_1 \\ \vdots \\ y_k \end{pmatrix} \quad U_1 = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_k \end{pmatrix} \quad Y_2 = \begin{pmatrix} y_{k+1} \\ \vdots \\ y_p \end{pmatrix} \quad U_2 = \begin{pmatrix} \mu_{k+1} \\ \vdots \\ \mu_p \end{pmatrix}.$$

Note that R_{11} and V_{11} are $k \times k$.

The following theorem (Theorem 9.11 on page 213) is critical for what follows:

The conditional distribution of Y_1 given Y_2 is the *k*-variate normal with mean

$$U_1 + V_{12}V_{22}^{-1}(Y_2 - U_2)$$

and covariance matrix

$$\mathrm{R}_{11}^{-1} = V_{11} - V_{12}V_{22}^{-1}V_{21} \quad .$$

Note that the covariance matrix of Y_1 given Y_2 does not depend on what the value of Y_2 is. This fact will be very important.

A ***probabilistic scenario*** arises when, instead of forecasting a single realization for the vector Y_2, the user forecasts a probability distribution for the vector Y_2. There is, however, a consistency issue because the true marginal distribution of Y_2 is given by

$$g(Y_2) = \frac{|V_{22}|^{-\frac{1}{2}}}{(2\pi)^{\frac{(p-k)}{2}}} e^{-\frac{1}{2}(Y_2-U_2)' V_{22}^{-1}(Y_2-U_2)} \quad .$$

Also, by definition, the conditional distribution of $y_1, y_2, \ldots, y_k$ given $y_{k+1}, y_{k+2}, \ldots, y_p$ is

$$f(Y_1 | Y_2) = \frac{h(\mathbf{Y})}{g(Y_2)} \quad .$$

Therefore, any forecast by the user other than the probability distribution $g(Y_2)$ is inconsistent with the underlying joint probability distribution $h(\mathbf{Y})$ stated above.

Nevertheless, an economic forecast often entails the belief, perhaps mistaken, that special information can be used to infer that the future will be different from the past.

3. The user-supplied distribution for a probabilistic scenario

Some additional notation is required. As above,

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \quad \text{where} \quad Y_1 = \begin{pmatrix} y_1 \\ \vdots \\ y_k \end{pmatrix} \quad \text{and} \quad Y_2 = \begin{pmatrix} y_{k+1} \\ \vdots \\ y_p \end{pmatrix} .$$

A particular realization of the vector Y_2, the i th, will be denoted by

$$y_2^i \equiv \begin{pmatrix} y_{k+1}(i) \\ \vdots \\ y_p(i) \end{pmatrix} .$$

We take as given a user forecast for the vectors y_2^i and their associated probabilities p_i. However this user forecast is generated, we take as given the following (marginal) distribution of Y_2:

$$\Pr(Y_2 = y_2^i) \equiv \Pr\left[Y_2 = \begin{pmatrix} y_{k+1}(i) \\ \vdots \\ y_p(i) \end{pmatrix} \right] = p_i, \quad p_i \geq 0, \ \sum_i p_i = 1 \quad . \tag{1}$$

This discrete distribution will be denoted by $\xi(Y_2)$.

4. The probabilistic scenario expected value and variance

Given the user forecast, the expected value of Y_2 is

$$E(Y_2;\xi) = \sum_i p_i y_2^i \equiv U_2^C \tag{2}$$

where the notation $E(Y_2;\xi)$ denotes that the expectation is taken with respect to the user-provided distribution $\xi(Y_2)$ and where the superscript *C* denotes "conditional on the user forecast."

Similarly, the $(p-k)\times(p-k)$ covariance matrix of Y_2 is given by

$$\begin{aligned}\operatorname{cov}(Y_2;\xi) &= E\left\{\left[Y_2 - E(Y_2)\right]\left[Y_2 - E(Y_2)\right]' ;\xi\right\} \\ &= \sum_i p_i \left\{\left[y_2^i - \sum_i p_i y_2^i\right]\left[y_2^i - \sum_i p_i y_2^i\right]'\right\} \\ &\equiv V_{22}^C .\end{aligned} \tag{3}$$

Equations (2) and (3) follow directly from the definitions of expected value and covariance.

We require the following Sliced Normal Theorem to infer anything about the distribution of Y_1 given the user forecast $\xi(Y_2)$. Note also that the user forecast alone tells us nothing about the covariance of Y_1 and Y_2.

5. The Sliced Normal Theorem

From the well-known results stated above in Section 2, for each realization of the random vector $Y_2 = y_2^i$, the conditional distribution $f\left(Y_1 | Y_2 = y_2^i\right)$ is multivariate normal and has a known mean and covariance matrix:

$$\left(Y_1 | y_2^i\right) \sim \mathscr{N}\left[U_1 + V_{12}V_{22}^{-1}\left(y_2^i - U_2\right), V_{11} - V_{12}V_{22}^{-1}V_{21}\right] . \tag{4}$$

For each *i* these conditional distributions are slices of the multivariate normal distribution, scaled so that the volume under the density function is one. That is, using the results from Section 2,

$$f\left(Y_1 | y_2^i\right) = \frac{h\left(Y_1, y_2^i\right)}{g\left(y_2^i\right)}$$

where $\dfrac{1}{g\left(y_2^i\right)}$ is the scale factor that makes the volume one. Here $g(Y_2)$ is the *true* marginal distribution of Y_2, *not* the user forecast $\xi(Y_2)$.

The sliced normal distribution for the $p\times 1$ vector **Y** is defined by these slices and the user-supplied distribution $\xi(Y_2)$ with mean $U_2^C \equiv \sum_i p_i y_2^i$ and covariance matrix V_{22}^C. Note that **Y** is not multivariate normal. By Theorem 9.11 of Mood and Graybill stated above, Y_1 is a mixture of multivariate normals, while the discrete distribution for Y_2 is $\xi(Y_2)$ defined above. Therefore the joint distribution of $\mathbf{Y} = (Y_1, Y_2)$ is complicated.

More formally, from (4) we know that the conditional density of the $k\times 1$ vector Y_1 for each given $(p-k)\times 1$ vector $Y_2 = y_2^i$ is

$$f\left(Y_1 \middle| Y_2 = y_2^i\right) = \frac{1}{(2\pi)^{\frac{k}{2}}\sqrt{\left|V_{11} - V_{12}V_{22}^{-1}V_{21}\right|}} e^{-\frac{1}{2}\left[Y_1 - \left(U_1 + V_{12}V_{22}^{-1}(y_2^i - U_2)\right)\right]'\left[V_{11} - V_{12}V_{22}^{-1}V_{21}\right]^{-1}\left[Y_1 - \left(U_1 + V_{12}V_{22}^{-1}(y_2^i - U_2)\right)\right]}. \quad (5)$$

Multiplying (5) by the probability p_i gives the equation for the i th slice of the sliced joint normal:

$$\varphi(Y_1, y_2^i) \equiv p_i f\left(Y_1 \middle| y_2^i\right) \quad \text{for} \quad -\infty < y_j < \infty,\ j = 1,\ldots,k;\ Y_2 = y_2^i .$$

Hence, given the user's forecast for the n realizations $y_2^i,\ i = 1, 2, \ldots, n$, the joint probability density function for the p-variate sliced normal is

$$\varphi(Y_1, Y_2) \equiv \begin{cases} p_1 f\left(Y_1 \middle| y_2^1\right) & \text{for} \quad -\infty < y_j < \infty,\ j = 1, \ldots, k;\ Y_2 = y_2^1 \\ & \vdots \\ p_i f\left(Y_1 \middle| y_2^i\right) & \text{for} \quad -\infty < y_j < \infty,\ j = 1, \ldots, k;\ Y_2 = y_2^i \\ & \vdots \\ p_n f\left(Y_1 \middle| y_2^n\right) & \text{for} \quad -\infty < y_j < \infty,\ j = 1, \ldots, k;\ Y_2 = y_2^n \end{cases} . \quad (6)$$

Of course, the conditional distribution of Y_1 given Y_2 for the sliced normal distribution is $f\left(Y_1 \middle| Y_2 = y_2^i\right)$, while the marginal distribution of Y_2 is $\Pr\left(Y_2 = y_2^i\right) = p_i$. The marginal distribution of Y_1 is

$$\sum_i p_i\, f\left(Y_1 | y_2^i\right) \quad \text{for} \quad -\infty < y_j < \infty,\ j = 1, \ldots, k;\ \ Y_2 = y_2^i,\ i = 1, \ldots, n. \ .$$

The expected value of the vector Y_1 for the sliced normal distribution is

$$
\begin{aligned}
U_1^C &\equiv E(Y_1) \\
&= \sum_i \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p_i Y_1 f(Y_1 | y_2^i) dy_1 \cdots dy_k \\
&= \sum_i p_i \left[E\left(Y_1 | y_2^i\right)\right] \\
&= \sum_i p_i \left[U_1 + V_{12} V_{22}^{-1} \left(y_2^i - U_2 \right)\right] = U_1 + V_{12} V_{22}^{-1} \sum_i p_i \left(y_2^i - U_2 \right) \\
&= U_1 + V_{12} V_{22}^{-1} \left(U_2^C - U_2 \right) .
\end{aligned} \tag{7}
$$

Similarly the expected value of the vector Y_2 for the sliced normal distribution is

$$
\begin{aligned}
U_2^C &\equiv E(Y_2) \\
&= \sum_i \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p_i y_2^i f(Y_1 | y_2^i) dy_1 \cdots dy_k \\
&= \sum_i p_i y_2^i \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(Y_1 | y_2^i) dy_1 \cdots dy_k \\
&= \sum_i p_i y_2^i \cdot (1) \\
&= \sum_i p_i y_2^i .
\end{aligned} \tag{8}
$$

We now state the formal result:

Sliced Normal Theorem

The p-variate Sliced Normal Distribution defined by (6) has mean

$$
\begin{bmatrix} U_1^C \\ U_2^C \end{bmatrix} = \begin{bmatrix} U_1 + V_{12} V_{22}^{-1} \left(U_2^C - U_2 \right) \\ \sum_i p_i y_2^i \end{bmatrix} \tag{9}
$$

and covariance matrix

$$
V^* = \begin{bmatrix} V_{11} - V_{12} V_{22}^{-1} \left(V_{22} - V_{22}^C \right) V_{22}^{-1} V_{21} & V_{12} V_{22}^{-1} V_{22}^C \\ V_{22}^C V_{22}^{-1} V_{21} & V_{22}^C \end{bmatrix} . \tag{10}
$$

Equations (7) and (8) establish (9). We now must prove (10). Note that $V_{22}^* = V_{22}^C$ by definition. The difficulty involves computing the covariance matrices $V_{12}^* = \left(V_{21}^*\right)'$ and V_{11}^*.

6. Computation of V_{11}^*

By definition

$$V_{11}^* = \sum_i \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p_i \left(Y_1 - U_1^C\right)\left(Y_1 - U_1^C\right)' f\left(Y_1 \middle| y_2^i\right) dy_1 \cdots dy_k \quad . \tag{11}$$

To ease notation we define

$$b \equiv E\left(Y_1 \middle| y_2^i\right) = U_1 + V_{12}V_{22}^{-1}\left(y_2^i - U_2\right) \ ; \tag{12}$$

see equation (7). Then (11) may be written as

$$V_{11}^* =$$
$$\sum_i \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p_i \left\{\left[Y_1 - E\left(Y_1 \middle| y_2^i\right)\right] + \left[E\left(Y_1 \middle| y_2^i\right) - U_1^C\right]\right\}\left\{\left[Y_1 - E\left(Y_1 \middle| y_2^i\right)\right] + \left[E\left(Y_1 \middle| y_2^i\right) - U_1^C\right]\right\}' f\left(Y_1 \middle| y_2^i\right) dy_1 \cdots dy_k =$$
$$\sum_i \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p_i \left[\left(Y_1 - b\right) + \left(b - U_1^C\right)\right]\left[\left(Y_1 - b\right) + \left(b - U_1^C\right)\right]' f\left(Y_1 \middle| y_2^i\right) dy_1 \cdots dy_k \ .$$

Performing the multiplication in the integrand gives

$$\begin{aligned} \sum_i \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p_i \Big[& \left(Y_1 - b\right)\left(Y_1 - b\right)' + \left(b - U_1^C\right)\left(Y_1 - b\right)' \\ & + \left(Y_1 - b\right)\left(b - U_1^C\right)' + \left(b - U_1^C\right)\left(b - U_1^C\right)' \Big] f\left(Y_1 \middle| y_2^i\right) dy_1 \cdots dy_k \ . \end{aligned} \tag{13}$$

We now proceed to evaluate the four terms in (13).

$$\begin{aligned} & \sum_i \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p_i \left(Y_1 - b\right)\left(Y_1 - b\right)' f\left(Y_1 \middle| y_2^i\right) dy_1 \cdots dy_k = \\ & \sum_i p_i \left[\operatorname{var}\left(Y_1 \middle| y_2^i\right)\right] = \sum_i p_i \left(V_{11} - V_{12}V_{22}^{-1}V_{21}\right) = V_{11} - V_{12}V_{22}^{-1}V_{21} \ . \end{aligned} \tag{14}$$

$$
\begin{aligned}
&\sum_i \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} p_i\left(b-U_1^C\right)\left(Y_1-b\right)' f\left(Y_1\middle|y_2^i\right)dy_1\cdots dy_k = \\
&\sum_i \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} p_i\left(Y_1-b\right)\left(b-U_1^C\right)' f\left(Y_1\middle|y_2^i\right)dy_1\cdots dy_k = \\
&\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\left[Y_1-E\left(Y_1\middle|y_2^i\right)\right]f\left(Y_1\middle|y_2^i\right)dy_1\cdots dy_k \sum_i p_i\left(b-U_1^C\right)' = \\
&(0)\cdot\sum_i p_i\left(b-U_1^C\right)' = 0 \ .
\end{aligned} \tag{15}
$$

We will use the following result:

$$
\begin{aligned}
\left(b-U_1^C\right) &= \left[U_1+V_{12}V_{22}^{-1}\left(y_2^i-U_2\right)\right]-\left[U_1+V_{12}V_{22}^{-1}\left(U_2^C-U_2\right)\right] \\
&= V_{12}V_{22}^{-1}\left(y_2^i-U_2^C\right) \ .
\end{aligned} \tag{16}
$$

$$
\begin{aligned}
&\sum_i \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} p_i\left(b-U_1^C\right)\left(b-U_1^C\right)' f\left(Y_1\middle|y_2^i\right)dy_1\cdots dy_k = \quad \text{using (16)} \\
&\sum_i \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} p_i\left[V_{12}V_{22}^{-1}\left(y_2^i-U_2^C\right)\right]\left[V_{12}V_{22}^{-1}\left(y_2^i-U_2^C\right)\right]' f\left(Y_1\middle|y_2^i\right)dy_1\cdots dy_k = \\
&\sum_i \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} p_i V_{12}V_{22}^{-1}\left(y_2^i-U_2^C\right)\left(y_2^i-U_2^C\right)' V_{22}^{-1}V_{21} f\left(Y_1\middle|y_2^i\right)dy_1\cdots dy_k = \\
&\sum_i p_i\left[V_{12}V_{22}^{-1}\left(y_2^i-U_2^C\right)\left(y_2^i-U_2^C\right)' V_{22}^{-1}V_{21}\right]\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f\left(Y_1\middle|y_2^i\right)dy_1\cdots dy_k = \\
&\sum_i p_i\left[V_{12}V_{22}^{-1}\left(y_2^i-U_2^C\right)\left(y_2^i-U_2^C\right)' V_{22}^{-1}V_{21}\right]\cdot 1 = \\
&V_{12}V_{22}^{-1}V_{22}^C V_{22}^{-1}V_{21} \ .
\end{aligned} \tag{17}
$$

Substituting (14), (15), and (17) into (13) gives

$$
\begin{aligned}
V_{11}^* &= V_{11}-V_{12}V_{22}^{-1}V_{21}+V_{12}V_{22}^{-1}V_{22}^C V_{22}^{-1}V_{21} \\
&= V_{11}-V_{12}V_{22}^{-1}\left(V_{22}-V_{22}^C\right)V_{22}^{-1}V_{21} \ .
\end{aligned} \tag{18}
$$

7. Computation of V_{21}^*

By definition

$$V_{21}^* = \sum_i \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p_i \left(y_2^i - U_2^C\right)\left(Y_1 - U_1^C\right)' f\left(Y_1 \middle| y_2^i\right) dy_1 \cdots dy_k$$

$$= \sum_i p_i \left(y_2^i - U_2^C\right)\left[E\left(Y_1 \middle| y_2^i\right) - U_1^C\right]'$$

$$= \text{ using (12) and (16) } \sum_i p_i \left(y_2^i - U_2^C\right)\left[V_{12} V_{22}^{-1}\left(y_2^i - U_2^C\right)\right]' \tag{19}$$

$$= \sum_i p_i \left(y_2^i - U_2^C\right)\left(y_2^i - U_2^C\right)' V_{22}^{-1} V_{21}$$

$$= V_{22}^C V_{22}^{-1} V_{21} \ .$$

Then

$$V_{12}^* = \left(V_{21}^*\right)'$$
$$= V_{12} V_{22}^{-1} V_{22}^C \ . \tag{20}$$

Comparison of (18), (19), and (20) with (10) completes the proof of the Sliced Normal Theorem.

8. Relationship to ordinary least squares

As is well-known, much of the above is closely related to ordinary least squares regression. To illustrate this fact, consider the multiple regression equation

$$Y_1 = U_1 + \beta\left(Y_2 - U_2\right) + \varepsilon \tag{21}$$

where $\beta \equiv V_{12} V_{22}^{-1}$ is a $k \times (p-k)$ matrix. Taking conditional expectations of (21) gives

$$E\left(Y_1 \middle| Y_2\right) = U_1 + \beta\left(Y_2 - U_2\right) . \tag{22}$$

Moreover, the covariance matrix of the error term is

$$
\begin{aligned}
E(\varepsilon\varepsilon') &= E\left\{\left[Y_1 - \left(U_1 + \beta(Y_2 - U_2)\right)\right]\left[Y_1 - \left(U_1 + \beta(Y_2 - U_2)\right)\right]'\right\} \\
&= E\left[(Y_1 - U_1)(Y_1 - U_1)'\right] + E\left[\beta(Y_2 - U_2)(Y_2 - U_2)'\beta'\right] - \\
&\quad E\left[(Y_1 - U_1)(Y_2 - U_2)'\beta'\right] - E\left[\beta(Y_2 - U_2)(Y_1 - U_1)'\right] \\
&= V_{11} + \beta V_{22}\beta' - V_{12}\beta' - V_{21}\beta' \\
&= V_{11} + V_{12}V_{22}^{-1}V_{21} - V_{12}V_{22}^{-1}V_{21} - V_{12}V_{22}^{-1}V_{21} \\
&= V_{11} - V_{12}V_{22}^{-1}V_{21}
\end{aligned} \tag{23}
$$

which is $\operatorname{cov}(Y_1 | Y_2)$. Also

$$
\begin{aligned}
E\left[(Y_2 - U_2)\varepsilon'\right] &= E\left\{(Y_2 - U_2)\left[Y_1 - \left(U_1 + \beta(Y_2 - U_2)\right)\right]'\right\} \\
&= E\left[(Y_2 - U_2)(Y_1 - U_1)\right] - E\left[(Y_2 - U_2)(Y_2 - U_2)'\beta'\right] \\
&= V_{21} - V_{22}V_{22}^{-1}V_{21} \\
&= 0
\end{aligned} \tag{24}
$$

so the error term is orthogonal to the right-hand-side variables.

*9. Positive definiteness of the covariance matrix V**

Of course, the covariance matrix V^* for the sliced normal distribution given by (10) must be positive definite by definition. Nevertheless, it is an instructive check to verify this fact directly.

We will use the following results:

Let the $p \times p$ *matrix* $\mathbf{A}$ *be partioned into* $\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ *where* A_{11} *is* $k \times k$, A_{12} *is* $k \times (p-k)$, A_{21} *is* $(p-k) \times k$, *and* A_{22} *is* $(p-k) \times (p-k)$.

Then $\begin{vmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{vmatrix} = \begin{vmatrix} I & 0 \\ 0 & A_{22} \end{vmatrix} \cdot \begin{vmatrix} A_{11} & A_{12} \\ 0 & I \end{vmatrix} = |A_{11}| \cdot |A_{22}|$. (25)

A reference for (25) is T. W. Anderson, ***An Introduction to Multivariate Statistical Analysis***, 2nd Edition, New York: John Wiley & Sons, 1984, pages 592-593.

If **P** *is a nonsingular matrix and if* **A** *is positive definite (semidefinite), then* **P′AP** *is positive definite (semidefinite).* (26)

A reference for (26) is Henri Theil, ***Principles of Econometrics***, New York: John Wiley & Sons, 1971, Proposition F3, page 22.

We now prove:

The $k \times k$ *matrix* $\left[V_{11} - V_{12}V_{22}^{-1}V_{21}\right]$ *is positive definite.* (27)

Proof:

Let x be any $k \times 1$ vector, $x \neq 0$. Then

$$\begin{pmatrix} x' & -x'V_{12}V_{22}^{-1} \end{pmatrix} \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} \begin{bmatrix} x \\ -V_{22}^{-1}V_{21}x \end{bmatrix} =$$

$$\begin{bmatrix} x'V_{11} - x'V_{12}V_{22}^{-1}V_{21} & x'V_{12} - x'V_{12}V_{22}^{-1}V_{22} \end{bmatrix} \begin{bmatrix} x \\ -V_{22}^{-1}V_{21}x \end{bmatrix} =$$

$$\begin{bmatrix} x'\left(V_{11} - V_{12}V_{22}^{-1}V_{21}\right) & x' \cdot (0) \end{bmatrix} \begin{bmatrix} x \\ -V_{22}^{-1}V_{21}x \end{bmatrix} =$$

$$x'\left(V_{11} - V_{12}V_{22}^{-1}V_{21}\right)x > 0$$

because V is positive definite.

Proposition

V* is positive definite.

Proof:

$$Let\ \ P' = \begin{bmatrix} I & -V_{12}V_{22}^{-1} \\ 0 & I \end{bmatrix},\ \ P = \begin{bmatrix} I & 0 \\ -V_{22}^{-1}V_{21} & I \end{bmatrix}.$$

Then

$$P'V^*P = \begin{bmatrix} V_{11} - V_{12}V_{22}^{-1}\left(V_{22} - V_{22}^C\right)V_{22}^{-1}V_{21} - V_{12}V_{22}^{-1}V_{22}^C V_{22}^{-1}V_{21} & V_{12}V_{22}^{-1}V_{22}^C - V_{12}V_{22}^{-1}V_{22}^C \\ V_{22}^C V_{22}^{-1}V_{21} & V_{22}^C \end{bmatrix} P$$

$$= \begin{bmatrix} V_{11} - V_{12}V_{22}^{-1}V_{21} & 0 \\ V_{22}^C V_{22}^{-1}V_{21} & V_{22}^C \end{bmatrix} P$$

$$= \begin{bmatrix} V_{11} - V_{12}V_{22}^{-1}V_{21} & 0 \\ V_{22}^C V_{22}^{-1}V_{21} - V_{22}^C V_{22}^{-1}V_{21} & V_{22}^C \end{bmatrix}$$

$$= \begin{bmatrix} V_{11} - V_{12}V_{22}^{-1}V_{21} & 0 \\ 0 & V_{22}^C \end{bmatrix} \equiv V^{**}.$$

Moreover, by property (25) of determinants stated above, $|P| = |P'| = 1$*. Hence* P^{-1} *and* $(P')^{-1}$ *both exist, and we may write* $V^* = (P')^{-1} V^{**} P^{-1}$*. Therefore, using (25), (26), and (27) above and the fact that* V_{22}^C *is positive definite by construction, we conclude that* V^* *is positive definite.*

Alternative proof:

For any $p \times 1$ vector $x \neq 0$, let $y = P^{-1}x$. Then

$x'V^*x =$

$$= x'\left[(P')^{-1}P'\right]V^*\left[PP^{-1}\right]x = x'(P')^{-1}\left(P'V^*P\right)P^{-1}x$$

$$= y'V^{**}y > 0$$

because V^{**} is postive definite.

Corollary

Suppose the forecasted covariance matrix is given by

$$V_{22}^C = \alpha' V_{22} \alpha$$

where α is a $(p-k)\times(p-k)$ diagonal matrix with positive diagonal elements. Then the forecasted standard deviation of the variable $k+i$ is then equal to α_{k+i} times its historical standard deviation, while the forecasted correlations between any two forecasted variables are the same as their historical correlations. The covariance matrix V^* is positive definite in this case.

If some diagonal elements of the matrix α are allowed to be zero (a point forecast with probability one), then V^* is positive semi-definite.