Mean-Variance Analysis and the Diversification of Risk

by Leigh H. Halliwell

<u>Mean-Variance Analysis and the Diversification of Risk</u>

Leigh J. Halliwell

ABSTRACT

Harry W. Markowitz in the 1950's developed mean-variance analysis, the theory of combining risky assets so as to minimize the variance of return (i.e., risk) at any desired mean return. The locus of optimal mean-variance combinations is called the efficient frontier, on which all rational investors desire to be positioned.

Actuaries see diagrams of efficient frontiers in their finance readings. Perhaps they are aware that efficient frontiers are parabolic. However, no mathematics is ever presented, so actuaries would be at a loss to derive an efficient frontier for problems involving more than two assets. But the minimum-variance combination of assets as a function of expected return has a simple matrix formulation; and the derivation of this formula is well within the grasp of actuaries. From this follows the formula for the efficient frontier.

This paper will present the mathematical theory of the efficient frontier. Then the theory will be illustrated by deriving the efficient frontier of a portfolio of stocks, treasury bonds, and treasury bills, as discussed in Ibbotson's *Stocks, Bonds, Bills, and Inflation 1994 Yearbook*. Also shown will be how to determine the mix of annual statement items which minimizes risk-based capital. An appendix will delve into the theory more deeply.

Mr. Halliwell is an Associate of the Casualty Actuarial Society and a member of the American Academy of Actuaries. Since April of 1993 he has been the Chief Actuary of the Louisiana Workers' Compensation Corporation in Baton Rouge, LA. Prior to that he worked at the National Council on Compensation Insurance in Boca Raton, FL.

1) PORTFOLIOS AS MATRICES

We have a portfolio of n assets, the return of the i^{th} asset, R_i, being a random variable with mean μ_i. We will let **R** denote the (n x 1) vector whose i^{th} element is R_i. Similarly, **M** will denote the (n x 1) vector whose i^{th} element is $\mu_i = E(R_i)$. Let Σ denote the symmetric (n x n) matrix whose $(i, j)^{th}$ element is $\sigma_{ij} = Cov(R_i, R_j)$. In matrix terms, $\mathbf{M} = E(\mathbf{R})$, and $\Sigma = Var(\mathbf{R}) = E((\mathbf{R}-\mathbf{M})(\mathbf{R}-\mathbf{M})')$, where ' is the symbol for transposition.[1] We will write $\mathbf{R} \sim [\mathbf{M}, \Sigma]$ as shorthand for '**R** is a random vector with mean **M** and variance Σ.' We will not make any assumptions as to the probability distribution of **R**; in particular, **R** need not be multivariate normal.

Our portfolio contains the assets in some proportion. Let Ω denote the (n x 1) vector whose elements represent the weights of the assets according to their market values; in other words, Ω represents the allocation of the portfolio's market value among its components. For Ω to be a true allocation vector its elements must sum to unity, or $\mathbf{J}_n'\Omega = \mathbf{1}$, where $\mathbf{J}_n$ is an (n x 1) vector of ones. The elements of Ω are real numbers, whether positive or negative. Negative weights are feasible in that assets may be borrowed or sold short.

The portfolio $(\mathbf{R}, \Omega)$ has overall return characteristics: its mean is $E(\Omega'\mathbf{R}) = \Omega'\mathbf{M}$, and its variance is $Var(\Omega'\mathbf{R}) = E(\{\Omega'\mathbf{R} - \Omega'M\}\{\Omega'\mathbf{R} - \Omega'M\}') = E(\Omega'\{\mathbf{R} - M\}\{\mathbf{R} - M\}'\Omega) = \Omega'E(\{\mathbf{R} - M\}\{\mathbf{R} - M\}')\Omega = \Omega'\Sigma\Omega$. Therefore, $\Omega'\mathbf{R} \sim [\Omega'\mathbf{M}, \Omega'\Sigma\Omega]$. Since the (1 x 1) matrix $\Omega'\Sigma\Omega$ is a variance, it must be greater than or equal to $\mathbf{0}$, irrespective of Ω. A symmetric matrix Σ such that $\Omega'\Sigma\Omega$ is nonnegative for all Ω is said to be *nonnegative definite*. Every real-valued random vector with finite covariances must have a nonnegative definite variance. If Σ has the additional property that $\Omega'\Sigma\Omega = \mathbf{0}$ only if $\Omega = \mathbf{0}$, then Σ is said to be *positive definite*. If the variance of a random vector is positive definite, then every non-zero linear combination of its elements has a positive variance. Therefore, if the variance of a group of assets is positive definite, then no allocation among them will be risk-free.

2) THE EFFICIENT FRONTIER

In general, without a utility function $U(\mu, \sigma^2)$ we cannot compare and rank various mean-variance combinations. However, among combinations having the same μ, we would assign the greatest utility to combinations having the smallest σ^2, which combinations by definition lie on the efficient frontier. In the context of our portfolio, we want to identify the Ω's which yield optimal mean-variance characteristics, i.e., points on the efficient frontier. We have our return vector $\mathbf{R} \sim [\mathbf{M}, \Sigma]$, and we will assume that Σ is positive definite. So in our portfolio there is neither a riskless asset, nor a riskless combination of assets. It is a theorem of nonnegative definite matrices that they are also positive definite if and only if they can be inverted. Σ is assumed to be positive definite, because in what follows Σ^{-1} must exist. Another theorem is that the inverse of a positive definite matrix is itself positive definite.[2]

First, let us find an allocation Ω_0 which yields the absolutely smallest $\sigma_0^2 = \Omega_0'\Sigma\Omega_0$. One such allocation is $\Omega_0 = \Sigma^{-1} \mathbf{J}_n (\mathbf{J}_n'\Sigma^{-1} \mathbf{J}_n)^{-1}$. Ω_0 is qualified because it exists and is an allocation vector. Its existence depends on the existence of the (1 x 1) matrix $(\mathbf{J}_n'\Sigma^{-1} \mathbf{J}_n)^{-1}$, which in turn exists if and only if $\mathbf{J}_n'\Sigma^{-1} \mathbf{J}_n$ is non-zero. But this is certain because Σ^{-1} is positive definite, according to a theorem stated in the previous paragraph. Furthermore, Ω_0 is an allocation vector because $\mathbf{J}_n'\Omega_0 = \mathbf{J}_n'\Sigma^{-1} \mathbf{J}_n (\mathbf{J}_n'\Sigma^{-1} \mathbf{J}_n)^{-1} = \mathbf{1}$.

The associated variance $\sigma_0^2 = \Omega_0'\Sigma\Omega_0 = \{(\mathbf{J}_n'\Sigma^{-1} \mathbf{J}_n)^{-1} \mathbf{J}_n'\Sigma^{-1}\}\Sigma\, \Sigma^{-1} \mathbf{J}_n (\mathbf{J}_n'\Sigma^{-1} \mathbf{J}_n)^{-1} = (\mathbf{J}_n'\Sigma^{-1} \mathbf{J}_n)^{-1} \mathbf{J}_n'\Sigma^{-1} \mathbf{J}_n (\mathbf{J}_n'\Sigma^{-1} \mathbf{J}_n)^{-1} = (\mathbf{J}_n'\Sigma^{-1} \mathbf{J}_n)^{-1}$. We can show that σ_0^2 is the absolute minimum by considering any other allocation vector $\Omega = \Omega_0 + \Delta$. Since Ω is an allocation vector, $1 = \mathbf{J}_n'\Omega = \mathbf{J}_n'\Omega_0 + \mathbf{J}_n'\Delta = \mathbf{1} + \mathbf{J}_n'\Delta$. Therefore, $\mathbf{J}_n'\Delta = \mathbf{0}$.

$$
\begin{aligned}
\sigma^2 = \Omega'\Sigma\Omega &= (\Omega_0 + \Delta)'\,\Sigma\,(\Omega_0 + \Delta) \\
&= \Omega_0'\,\Sigma\,\Omega_0 + \Omega_0'\,\Sigma\,\Delta + \Delta'\,\Sigma\,\Omega_0 + \Delta'\,\Sigma\,\Delta \\
&= \Omega_0'\,\Sigma\,\Omega_0 + 2\,\Omega_0'\,\Sigma\,\Delta + \Delta'\,\Sigma\,\Delta \\
&= \Omega_0'\,\Sigma\,\Omega_0 + 2\,\{\,\Sigma^{-1}\,\mathbf{J}_n\,(\mathbf{J}_n'\Sigma^{-1}\,\mathbf{J}_n)^{-1}\}'\,\Sigma\,\Delta + \Delta'\,\Sigma\,\Delta \\
&= \Omega_0'\,\Sigma\,\Omega_0 + 2\,\{(\mathbf{J}_n'\Sigma^{-1}\,\mathbf{J}_n)^{-1}\,\mathbf{J}_n'\Sigma^{-1}\}\Sigma\,\Delta + \Delta'\,\Sigma\,\Delta \\
&= \Omega_0'\,\Sigma\,\Omega_0 + 2\,(\mathbf{J}_n'\Sigma^{-1}\,\mathbf{J}_n)^{-1}\,\mathbf{J}_n'\Delta + \Delta'\,\Sigma\,\Delta \\
&= \Omega_0'\,\Sigma\,\Omega_0 + 2\,(\mathbf{J}_n'\Sigma^{-1}\,\mathbf{J}_n)^{-1}\,\mathbf{0} + \Delta'\,\Sigma\,\Delta \\
&= \Omega_0'\,\Sigma\,\Omega_0 + \Delta'\,\Sigma\,\Delta \\
&= \sigma_0^2 + \Delta'\,\Sigma\,\Delta
\end{aligned}
$$

Since Σ is positive definite, $\Delta'\,\Sigma\,\Delta$ is greater than or equal to $\mathbf{0}$, with equality holding if and only if $\Delta = \mathbf{0}$. Therefore, σ_0^2 is the absolute minimum variance, and Ω_0 is the only allocation vector which attains to it.

Next, for any μ, let us find an allocation $\Omega_0(\mu)$ which yields the smallest $\sigma_0^2(\mu) = \Omega_0(\mu)'\Sigma\Omega_0(\mu)$. $\Omega_0(\mu)$ differs from Ω_0 in that $\Omega_0(\mu)'\,\mathbf{M} = \mu$. Let us define an (n x 2) matrix $\mathbf{W}$ as the matrix whose first column is $\mathbf{J}_n$ and whose second column is $\mathbf{M}$. The constraints on $\Omega_0(\mu)$ may be expressed as:

$$\mathbf{W}'\Omega_0(\mu) = \begin{bmatrix} \mathbf{J}_n' \\ \mathbf{M}' \end{bmatrix} \Omega_0(\mu) = \begin{bmatrix} 1 \\ \mu \end{bmatrix}$$

In what follows we must assume that $\mathbf{W}$ is of full rank, i.e., that rank($\mathbf{W}$) = 2. According to one definition, the rank of a matrix is the largest dimension of a submatrix with a non-zero determinant. The rank of $\mathbf{W}$ must be less than min(n, 2), where $n > 1$. The rank will be two, if and only if there are at least two different elements of $\mathbf{M}$; otherwise it will be one. It is not restrictive to assume that at least two of the returns differ, and this assures that there will be a $\Omega_0(\mu)$ for each real-valued μ.[3] According to a theorem of matrix algebra, if rank($\mathbf{W}$) = 2, then the (2 x 2) matrix $\mathbf{W}'\Sigma^{-1}\,\mathbf{W}$ is positive definite, and hence invertible.

Let $\Omega_0(\mu) = \Sigma^{-1}\mathbf{W}(\mathbf{W}'\Sigma^{-1}\mathbf{W})^{-1}\begin{bmatrix}1\\ \mu\end{bmatrix}$. There is no problem with existence, since according to the assumptions of the previous paragraph $\mathbf{W}'\Sigma^{-1}\ \mathbf{W}$ has an inverse. Moreover, the constraints are satisfied, since $\mathbf{W}'\Omega_0(\mu) = \mathbf{W}'\Sigma^{-1}\mathbf{W}(\mathbf{W}'\Sigma^{-1}\mathbf{W})^{-1}\begin{bmatrix}1\\ \mu\end{bmatrix} = \begin{bmatrix}1\\ \mu\end{bmatrix}$. As for the variance:

$$
\begin{aligned}
\sigma_0^{\,2}(\mu) &= \Omega_0(\mu)'\Sigma^{-1}\Omega_0(\mu)\\
&= [1\quad \mu](\mathbf{W}'\Sigma^{-1}\mathbf{W})^{-1}\mathbf{W}'\Sigma^{-1}\Sigma\Sigma^{-1}\mathbf{W}(\mathbf{W}'\Sigma^{-1}\mathbf{W})^{-1}\begin{bmatrix}1\\ \mu\end{bmatrix}\\
&= [1\quad \mu](\mathbf{W}'\Sigma^{-1}\mathbf{W})^{-1}\mathbf{W}'\Sigma^{-1}\mathbf{W}(\mathbf{W}'\Sigma^{-1}\mathbf{W})^{-1}\begin{bmatrix}1\\ \mu\end{bmatrix}\\
&= [1\quad \mu](\mathbf{W}'\Sigma^{-1}\mathbf{W})^{-1}\begin{bmatrix}1\\ \mu\end{bmatrix}
\end{aligned}
$$

Similarly to the case of σ_0^2, we can show that $\sigma_0^2(\mu)$ is the minimum by considering any other allocation vector $\Omega(\mu) = \Omega_0(\mu) + \Delta$. Because of the constraint, $\mathbf{W}'\Omega(\mu) = \mathbf{W}'\Omega_0(\mu)$, so $\mathbf{W}'\Delta = \mathbf{0}$:

$$
\begin{aligned}
\sigma^2(\mu) &= \Omega(\mu)'\Sigma\Omega(\mu)\\
&= (\Omega_0(\mu) + \Delta)'\ \Sigma\ (\Omega_0(\mu) + \Delta)\\
&= \Omega_0(\mu)'\ \Sigma\ \Omega_0(\mu) + \Omega_0(\mu)'\ \Sigma\ \Delta + \Delta'\ \Sigma\ \Omega_0(\mu) + \Delta'\ \Sigma\ \Delta\\
&= \Omega_0(\mu)'\ \Sigma\ \Omega_0(\mu) + 2\ \Omega_0(\mu)'\ \Sigma\ \Delta + \Delta'\ \Sigma\ \Delta\\
&= \Omega_0(\mu)'\ \Sigma\ \Omega_0(\mu) + 2\ \{\ [1\ \mu]\ (\mathbf{W}'\Sigma^{-1}\mathbf{W})^{-1}\ \mathbf{W}'\Sigma^{-1}\}\Sigma\ \Delta + \Delta'\ \Sigma\ \Delta\\
&= \Omega_0(\mu)'\ \Sigma\ \Omega_0(\mu) + 2\ [1\ \mu]\ (\mathbf{W}'\Sigma^{-1}\mathbf{W})^{-1}\ \mathbf{W}'\ \Delta + \Delta'\ \Sigma\ \Delta\\
&= \Omega_0(\mu)'\ \Sigma\ \Omega_0(\mu) + 2\ [1\ \mu]\ (\mathbf{W}'\Sigma^{-1}\mathbf{W})^{-1}\ \mathbf{0} + \Delta'\ \Sigma\ \Delta\\
&= \Omega_0(\mu)'\ \Sigma\ \Omega_0(\mu) + \Delta'\ \Sigma\ \Delta\\
&= \sigma_0^2(\mu) + \Delta'\ \Sigma\ \Delta
\end{aligned}
$$

As before, since Σ is positive definite, $\Delta'\ \Sigma\ \Delta$ is greater than or equal to $\mathbf{0}$, with equality holding if and only if $\Delta = \mathbf{0}$. Therefore, $\sigma_0^2(\mu)$ is the minimum variance, and $\Omega_0(\mu)$ is the only constrained allocation vector which attains to it.

Notice that the minimum variance $\sigma_0^2(\mu) = \begin{bmatrix}1 & \mu\end{bmatrix}(\mathbf{W}'\Sigma^{-1}\mathbf{W})^{-1}\begin{bmatrix}1 \\ \mu\end{bmatrix}$ is quadratic in μ.

Therefore, there exist constants a, b, and c such that $\sigma_0^2(\mu) = a\,\mu^2 + b\,\mu + c$. If a were equal to zero, then $\sigma_0^2(\mu)$ would be linear in μ, which would imply that a unique absolute minimum does not exist. However, we saw above that an absolute minimum does exist, viz., $\sigma_0^2 = (\mathbf{J}_n'\Sigma^{-1}\,\mathbf{J}_n)^{-1}$. Therefore, a cannot be zero. Similarly, a cannot be negative. Hence, a must be positive, implying that the efficient frontier has the shape of a convex parabola.

3) AN ILLUSTRATION OF CALCULATING AN EFFICIENT FRONTIER

A discussion of the efficient frontier can be found in chapter 9 of *Stocks, Bonds, Bills, and Inflation 1994 Yearbook.*[4] There a portfolio is formed from three asset classes, large company stocks (the S&P 500), intermediate government bonds, and U.S. treasury bills. In that order, the expected return vector $\mathbf{M} = \begin{bmatrix}\mathbf{12.9\%} \\ \mathbf{5.3\%} \\ \mathbf{4.3\%}\end{bmatrix}$. The standard deviations we will represent as a (3 x 3) matrix, $\Lambda = \begin{bmatrix}\mathbf{20.5\%} & & \\ & \mathbf{6.5\%} & \\ & & \mathbf{2.8\%}\end{bmatrix}$; and the matrix of correlation coefficients $\mathrm{P} = \begin{bmatrix}\mathbf{100.0\%} & \mathbf{35.0\%} & \mathbf{-4.0\%} \\ \mathbf{35.0\%} & \mathbf{100.0\%} & \mathbf{16.0\%} \\ \mathbf{-4.0\%} & \mathbf{16.0\%} & \mathbf{100.0\%}\end{bmatrix}$. Our variance matrix, $\Sigma =$ $\Lambda\mathrm{P}\Lambda' = \begin{bmatrix}\mathbf{0.0420} & \mathbf{0.0047} & \mathbf{-0.0002} \\ \mathbf{0.0047} & \mathbf{0.0042} & \mathbf{0.0003} \\ \mathbf{-0.0002} & \mathbf{0.0003} & \mathbf{0.0008}\end{bmatrix}$. What is the efficient frontier of this portfolio?

At the vertex of the parabola:

$$\sigma_0^2 = (\mathbf{J}_3'\Sigma^{-1}\mathbf{J}_3)^{-1} = \left(\begin{bmatrix}1 & 1 & 1\end{bmatrix}\begin{bmatrix}\mathbf{0.0420} & \mathbf{0.0047} & \mathbf{-0.0002} \\ \mathbf{0.0047} & \mathbf{0.0042} & \mathbf{0.0003} \\ \mathbf{-0.0002} & \mathbf{0.0003} & \mathbf{0.0008}\end{bmatrix}\begin{bmatrix}1 \\ 1 \\ 1\end{bmatrix}\right)^{-1} = \begin{bmatrix}\mathbf{0.0007}\end{bmatrix}$$

$$\Omega_0 = \Sigma^{-1}\mathbf{J}_3(\mathbf{J}_3'\Sigma^{-1}\mathbf{J}_3)^{-1} = \Sigma^{-1}\mathbf{J}_3\sigma_0^2$$

$$= \begin{bmatrix} 0.0420 & 0.0047 & -0.0002 \\ 0.0047 & 0.0042 & 0.0003 \\ -0.0002 & 0.0003 & 0.0008 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} [0.0007]$$

$$= \begin{bmatrix} 1.1\% \\ 9.8\% \\ 89.1\% \end{bmatrix}$$

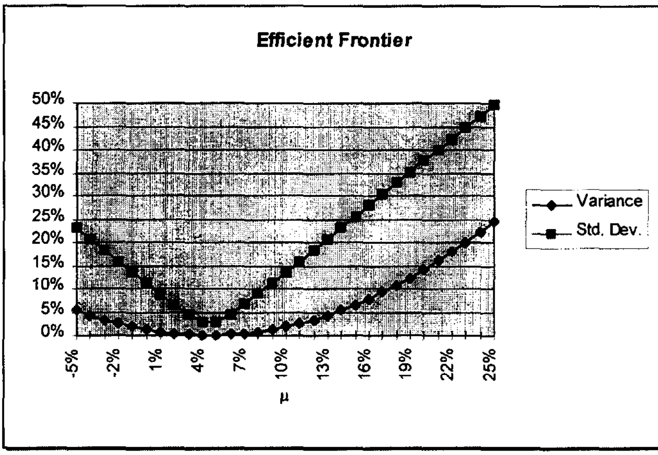
$$\mu_0 = \Omega_0'\mathbf{M} = [1.1\% \quad 9.8\% \quad 89.1\%] \begin{bmatrix} 12.9\% \\ 5.3\% \\ 4.3\% \end{bmatrix} = [4.5\%]$$

So we see that the closest thing to a risk-free portfolio is one with a variance of 0.0007, or with a standard deviation of 2.7%, which has an expected return of 4.5%.

For other points on the efficient frontier one must use the formula $\Omega_0(\mu) = \Sigma^{-1}\mathbf{W}(\mathbf{W}'\Sigma^{-1}\mathbf{W})^{-1}\begin{bmatrix} 1 \\ \mu \end{bmatrix}$, where $\mathbf{W} = \begin{bmatrix} 1 & 12.9\% \\ 1 & 5.3\% \\ 1 & 4.3\% \end{bmatrix}$. And one may calculate $\sigma_0^2(\mu) = [1 \quad \mu](\mathbf{W}'\Sigma^{-1}\mathbf{W})^{-1}\begin{bmatrix} 1 \\ \mu \end{bmatrix}$. A table of such values follows:

μ	$\sigma_0^2(\mu)$	Std. Dev.	$\Omega_0(\mu)$		
-5%	0.0534	23.1%	-1.1049	0.2019	1.9030
-4%	0.0429	20.7%	-0.9873	0.1909	1.7964
-3%	0.0336	18.3%	-0.8698	0.1799	1.6899
-2%	0.0254	15.9%	-0.7522	0.1689	1.5833
-1%	0.0184	13.6%	-0.6346	0.1579	1.4767
0%	0.0125	11.2%	-0.5171	0.1470	1.3701
1%	0.0079	8.9%	-0.3995	0.1360	1.2636
2%	0.0044	6.6%	-0.2820	0.1250	1.1570
3%	0.0020	4.5%	-0.1644	0.1140	1.0504
4%	0.0009	2.9%	-0.0469	0.1030	0.9438
5%	0.0009	3.0%	0.0707	0.0921	0.8373
6%	0.0020	4.5%	0.1882	0.0811	0.7307
7%	0.0044	6.6%	0.3058	0.0701	0.6241
8%	0.0079	8.9%	0.4234	0.0591	0.5175
9%	0.0126	11.2%	0.5409	0.0481	0.4109
10%	0.0184	13.6%	0.6585	0.0372	0.3044
11%	0.0255	16.0%	0.7760	0.0262	0.1978
12%	0.0337	18.3%	0.8936	0.0152	0.0912
13%	0.0430	20.7%	1.0111	0.0042	-0.0154
14%	0.0535	23.1%	1.1287	-0.0068	-0.1219
15%	0.0652	25.5%	1.2462	-0.0177	-0.2285
16%	0.0781	27.9%	1.3638	-0.0287	-0.3351
17%	0.0921	30.4%	1.4814	-0.0397	-0.4417
18%	0.1073	32.8%	1.5989	-0.0507	-0.5482
19%	0.1237	35.2%	1.7165	-0.0617	-0.6548
20%	0.1413	37.6%	1.8340	-0.0726	-0.7614
21%	0.1600	40.0%	1.9516	-0.0836	-0.8680
22%	0.1799	42.4%	2.0691	-0.0946	-0.9745
23%	0.2009	44.8%	2.1867	-0.1056	-1.0811
24%	0.2231	47.2%	2.3043	-0.1166	-1.1877
25%	0.2465	49.7%	2.4218	-0.1275	-1.2943

As expected, the higher the mean return, the more one must borrow, or short, the boı and the T-bills. A graph of the table follows:



This graph differs from those usually shown in finance texts in that the independent variable μ is graphed on the x-axis, rather than on the y-axis. Moreover, the parabola is the *variance* curve; most texts graph an efficient frontier of *standard deviation* versus mean.

4) OPTIMIZING A QUADRATIC FORM

The mean-variance optimization presented above is a special instance of the general problem of optimizing a quadratic form. A quadratric form is an expression of the form $\mathbf{X'\Sigma X}$, where $\mathbf{X}$ is (n x 1) and Σ is (n x n). Σ need not be symmetric; however, the quadratic form is symmetric since it is a (1 x 1) matrix. Therefore, $\mathbf{X'\Sigma X} = (\mathbf{X'\Sigma X})' = \mathbf{X'\Sigma' X}$. So $\mathbf{X'\Sigma X} = (1/2)\{\mathbf{X'\Sigma X} + \mathbf{X'\Sigma' X}\} = \mathbf{X'}\{(\Sigma + \Sigma')/2\}\mathbf{X}$, where $(\Sigma + \Sigma')/2$ is symmetric. So without loss of generality we may assume Σ to be symmetric.

The general problem is to find the $\mathbf{X}$ which minimizes $\mathbf{X'\Sigma X}$, subject to some constraint $\mathbf{AX} = \mathbf{B}$. Let us have m linear constraints, so $\mathbf{A}$ is (m x n) and $\mathbf{B}$ is (m x 1). If we choose the constraints so as to be independent of one another, then the rank of $\mathbf{A}$ will be m.

Provided that $(\mathbf{A}\Sigma^{-1}\mathbf{A}')^{-1}$ exists, the optimal $\mathbf{X}_0 = \Sigma^{-1}\mathbf{A}' (\mathbf{A}\Sigma^{-1}\mathbf{A}')^{-1} \mathbf{B}$.[5] The constraint is satisfied, since $\mathbf{A}\mathbf{X}_0 = \mathbf{A}\Sigma^{-1}\mathbf{A}' (\mathbf{A}\Sigma^{-1}\mathbf{A}')^{-1} \mathbf{B} = \mathbf{B}$. To show optimality, consider any $\mathbf{X} = \mathbf{X}_0 + \Delta$, such that $\mathbf{AX} = \mathbf{B}$. Therefore, $\mathrm{A}\,\Delta = \mathbf{0}$, and:

$$
\begin{aligned}
\mathbf{X}'\Sigma\mathbf{X} &= (\mathbf{X}_0 + \Delta)'\,\Sigma\,(\mathbf{X}_0 + \Delta) \\
&= \mathbf{X}_0'\,\Sigma\,\mathbf{X}_0 + \mathbf{X}_0'\,\Sigma\,\Delta + \Delta'\,\Sigma\,\mathbf{X}_0 + \Delta'\,\Sigma\,\Delta \\
&= \mathbf{X}_0'\,\Sigma\,\mathbf{X}_0 + 2\,\mathbf{X}_0'\,\Sigma\,\Delta + \Delta'\,\Sigma\,\Delta \\
&= \mathbf{X}_0'\,\Sigma\,\mathbf{X}_0 + 2\,\{\,\Sigma^{-1}\mathbf{A}'\,(\mathbf{A}\Sigma^{-1}\mathbf{A}')^{-1}\,\mathbf{B}\,\}'\,\Sigma\,\Delta + \Delta'\,\Sigma\,\Delta \\
&= \mathbf{X}_0'\,\Sigma\,\mathbf{X}_0 + 2\,\{\,\mathrm{B}'\,(\mathbf{A}\Sigma^{-1}\mathbf{A}')^{-1}\,\mathrm{A}\,\Sigma^{-1}\,\}\,\Sigma\,\Delta + \Delta'\,\Sigma\,\Delta \\
&= \mathbf{X}_0'\,\Sigma\,\mathbf{X}_0 + 2\,\mathrm{B}'\,(\mathbf{A}\Sigma^{-1}\mathbf{A}')^{-1}\,\mathbf{A}\,\Delta + \Delta'\,\Sigma\,\Delta \\
&= \mathbf{X}_0'\,\Sigma\,\mathbf{X}_0 + 2\,\mathrm{B}'\,(\mathbf{A}\Sigma^{-1}\mathbf{A}')^{-1}\,\mathbf{0} + \Delta'\,\Sigma\,\Delta \\
&= \mathbf{X}_0'\,\Sigma\,\mathbf{X}_0 + \Delta'\,\Sigma\,\Delta
\end{aligned}
$$

Therefore, if Σ is nonnegative definite, $\mathbf{X}_0'\,\Sigma\,\mathbf{X}_0$ is a minimum; moreover, if Σ is positive definite, it is a unique minimum. If Σ is not nonnegative definite, then $\mathbf{X}_0$ will be an inflection point, which in the context of quadratic forms is a saddle point. This generalized solution will be applied in the next section.

5) MINIMIZING RISK-BASED CAPITAL

Robert P. Butsic[6] presents an example of calculating the risk-based capital for a company whose balance sheet is:

Assets		Liabilities	
Stock	$ 200	Loss Reserve	$800
Bonds	1,000	Property UPR	100
Affiliates	100	Surplus	400

Butsic shows that the risk-based capital required for any basic element is approximately proportional to the standard deviation of a unit amount of that element [1, p. 343]. Since units of a basic element are perfectly correlated, it follows that the proper risk-based capital increases in proportion with the amount of the element. The basic elements of this example are stock, bonds, affiliated stock, loss reserve, and property UPR; and the proportionality constants expressed as a capital ratio (CR) vector are:

$$\mathbf{CR} = \begin{bmatrix} \textbf{Stock} \\ \textbf{Bonds} \\ \textbf{Affiliates} \\ \textbf{Loss\_Reserve} \\ \textbf{Property\_UPR} \end{bmatrix} = \begin{bmatrix} \mathbf{0.30} \\ \mathbf{0.05} \\ \mathbf{0.30} \\ \mathbf{0.40} \\ \mathbf{0.10} \end{bmatrix}$$

If we make CR a (5 x 5) diagonal matrix we can calculate the separate elements of risk-based capital:

$$\mathbf{CR*X} = \begin{bmatrix} \mathbf{0.30} & & & & \\ & \mathbf{0.05} & & & \\ & & \mathbf{0.30} & & \\ & & & \mathbf{0.40} & \\ & & & & \mathbf{0.10} \end{bmatrix} \begin{bmatrix} \mathbf{200} \\ \mathbf{1000} \\ \mathbf{100} \\ \mathbf{800} \\ \mathbf{100} \end{bmatrix} = \begin{bmatrix} \mathbf{60} \\ \mathbf{50} \\ \mathbf{30} \\ \mathbf{320} \\ \mathbf{10} \end{bmatrix}$$

However, we want to compute the risk-based capital of the surplus, and for this we need to know the covariances between the five elements. More precisely, since the capital ratios already incorporate the standard deviations of their elements, we need to know the correlation coefficients. Butsic provides us with the rho matrix:

$$\mathrm{P} = \begin{bmatrix} \mathbf{1.0} & \mathbf{0.2} & \mathbf{1.0} & & \\ \mathbf{0.2} & \mathbf{1.0} & \mathbf{0.2} & \mathbf{0.4} & \\ \mathbf{1.0} & \mathbf{0.2} & \mathbf{1.0} & \mathbf{-1.0} & \\ & \mathbf{0.4} & \mathbf{-1.0} & \mathbf{1.0} & \\ & & & & \mathbf{1.0} \end{bmatrix}$$

The formula, therefore, for the risk-based capital of any (5 x 1) combination of elements **X** is:

$$\begin{aligned} RBC^2 &= (CR*X)'P(CR*X) \\ &= X'(CR'*P*CR)X \\ &= X'(\Sigma)X \end{aligned}$$

where

$$\Sigma = \begin{bmatrix} 0.30 & & & & \\ & 0.05 & & & \\ & & 0.30 & & \\ & & & 0.40 & \\ & & & & 0.10 \end{bmatrix} \begin{bmatrix} 1.0 & 0.2 & 1.0 & & \\ 0.2 & 1.0 & 0.2 & 0.4 & \\ 1.0 & 0.2 & 1.0 & -1.0 & \\ & 0.4 & -1.0 & 1.0 & \\ & & & & 1.0 \end{bmatrix} \begin{bmatrix} 0.30 & & & & \\ & 0.05 & & & \\ & & 0.30 & & \\ & & & 0.40 & \\ & & & & 0.10 \end{bmatrix}$$

$$= \begin{bmatrix} 0.09 & 0.003 & 0.09 & & \\ 0.003 & 0.0025 & 0.003 & 0.008 & \\ 0.09 & 0.003 & 0.09 & -0.12 & \\ & 0.008 & -0.12 & 0.16 & \\ & & & & 0.01 \end{bmatrix}$$

Butsic has estimated the covariances as if the liabilities were negative assets; therefore, the \$400 surplus is represented by $X = \begin{bmatrix} 200 \\ 1000 \\ 100 \\ -800 \\ -100 \end{bmatrix}$. Therefore, $RBC^2 = \mathbf{X'\Sigma X} = 121{,}300$, or $RBC = \$348.28$.

But we might wish to know the **X** which minimizes $RBC^2 = \mathbf{X'\Sigma X}$, subject to some constraint $\mathbf{AX} = \mathbf{B}$. For example, we can attempt to minimize the risk-based capital of the surplus, given that the surplus is \$400, or $\mathbf{AX} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \end{bmatrix}\mathbf{X} = \begin{bmatrix} 400 \end{bmatrix}$. $(\mathbf{A\Sigma^{-1}A'})^{-1} = [0.001998]$, and

$\mathbf{X}_0 = \Sigma^{-1}\mathbf{A'}(\mathbf{A}\Sigma^{-1}\mathbf{A'})^{-1}\mathbf{B} = \begin{bmatrix} -17 \\ 322 \\ 15 \\ 0 \\ 80 \end{bmatrix}$. At this combination, $RBC^2 = \mathbf{X_0'\Sigma X_0} = 319.704$, or $RBC = \$17.88$.

Now suppose that we do not have so much freedom in apportioning our $400 of surplus. Rather, our only choice is how to apportion the first $1,200 between stock and bonds:

$$\mathbf{AX} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{X} = \begin{bmatrix} 1200 \\ 100 \\ -800 \\ -100 \end{bmatrix} = \mathbf{B}$$

In this case, $(\mathbf{A}\Sigma^{-1}\mathbf{A}')^{-1} = \begin{bmatrix} 0.002497 & 0.002497 & 0.008046 & 0 \\ 0.002497 & 0.002497 & -0.11195 & 0 \\ 0.008046 & -0.11195 & 0.15926 & 0 \\ 0 & 0 & 0 & 0.01 \end{bmatrix}$,

$\mathbf{X}_0 = \Sigma^{-1}\mathbf{A}'(\mathbf{A}\Sigma^{-1}\mathbf{A}')^{-1}\mathbf{B} = \begin{bmatrix} -182 \\ 1382 \\ 100 \\ -800 \\ -100 \end{bmatrix}$, and $\text{RBC}^2 = \mathbf{X}_0'\Sigma\mathbf{X}_0 = 108710$, or RBC = $329.71.

Many with good reason suppose that risk-based capital standards will induce insurers to sell some of their stocks; however, it surprises this author that risk-based capital might induce them to *short-sell* stocks. This may be only a peculiarity of Butsic's example; but the keepers of risk-based capital should test whether their parameters will lead to such undesirable results, if indeed this short-sale tendency be undesirable. Of course, short selling to minimize risk-based capital is countervailed by decreasing one's expected return on invested assets.

One more point needs to be made. There is a flaw in Butsic's example unrelated to the short-sale peculiarity just mentioned, viz., that the optimization assumed that Σ is a nonnegative definite matrix. However, it turns out that this particular Σ is not. A theorem states that a matrix is nonnegative definite if and only if all its eigenvalues are nonnegative. Here the eigenvalues of Σ are 0.2684, 0.1145, 0.0100, 0.0024, and -0.0428. This means that there are values of Δ for which $\Delta' \Sigma \Delta$ is negative, although they may not happen to be in the subspace in which $\mathbf{A}\,\Delta = \mathbf{0}$.[7] In any event, the negative eigenvalue originates in the correlation matrix P, and the keepers of risk-based capital should make sure that whatever P they use is at least nonnegative definite, and preferably positive definite.[8]

6) CONCLUSION

We have seen the power of mean-variance analysis, or more generally, quadratic-form optimization, in regard to risk diversification. One other application is to the problem of optimizing one's mix of business. Not having the requisite statistics, the author did not give an example of this. However, if one had mean operating ratios by line of business and a matrix of their covariances, then it would be easy to calculate an efficient frontier and the associated premium allocation vectors.

The essence of efficient diversification is to divide and conquer, i.e., to get one's adversaries to fight among themselves before they come against oneself. The author is reminded of a story in the book of Judges, according to which Gideon with an army of three hundred defeated a vastly superior Midianite army. The stated reason for this unlikely victory is that God had sent a spirit of confusion among the Midianites, so that they rose up "every man's sword against his fellow." Comedian George Carlin perhaps expressed the same idea in his bombastic suggestion for a fresh good-bye: "May the forces of evil become confused on the way to your house." In any event, actuaries should be employing mean-variance analysis in the construction of portofolios of risks.

NOTES

[1] It is presumed that the reader has some familiarity with matrix algebra. Therefore, some of the steps in the derivations may involve the application of multiple matrix theorems. Some of the basic properties of matricies are stated here, and may be of help if the reader is puzzled by a derivation:

A. Matrix multiplication is associative: A(BC)=(AB)C.
B. Matrix multiplication is not commutative; however, (1 x 1) matrices commute.
C. Matrix multiplication is distributive: A(B+C)=AB+AC.
D. Transposition of a product behaves thus: (AB)' = B'A'.
E. Similarly, with matrix inversion, $(AB)^{-1}=B^{-1}A^{-1}$, if A and B are nonsingular.
F. By definition, A is symmetric if and only if A'=A.
G. Every (1 x 1) matrix is symmetric.
H. If A is nonsingular, then $(A^{-1})^{-1} = A$. Also, $(A^{-1})'=(A')^{-1}$.

Judge [3] contains a seventy-five page appendix on matrix algebra that alone makes the book worth purchasing.

[2] Cf. Judge [3], pp. 960f., on definite matrices.

[3] A universe of assets all whose expected returns are the same is not realistic. However, the formula for the minimal variance $\sigma_0^2 = (\mathbf{J}_n'\Sigma^{-1}\mathbf{J}_n)^{-1}$ is still valid. There is a problem related to this situation, viz.: given $X_1, X_2, \ldots, X_n$, all of which have the same mean, what is the best linear unbiased estimator (BLUE) of that mean? If we may assume a matrix of covariances, Σ, the best estimator is the weighting of the X's according to $\Omega_0 = \Sigma^{-1}\mathbf{J}_n(\mathbf{J}_n'\Sigma^{-1}\mathbf{J}_n)^{-1}$. As a check, let us work through the familiar case in which the X's are uncorrelated, but have different variances:

$$\Sigma = \begin{bmatrix} \sigma_{11} & & & \\ & \sigma_{22} & & \\ & & \ddots & \\ & & & \sigma_{nn} \end{bmatrix}$$

$$\Sigma^{-1}J = \begin{bmatrix} \sigma_{11}^{-1} & & & \\ & \sigma_{22}^{-1} & & \\ & & \ddots & \\ & & & \sigma_{nn}^{-1} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} \sigma_{11}^{-1} \\ \sigma_{22}^{-1} \\ \vdots \\ \sigma_{nn}^{-1} \end{bmatrix}$$

$$J'\Sigma^{-1}J = \begin{bmatrix}1 & 1 & \cdots & 1\end{bmatrix}\begin{bmatrix}\sigma_{11}^{-1} \\ \sigma_{22}^{-1} \\ \vdots \\ \sigma_{nn}^{-1}\end{bmatrix} = \left[\sum_i \sigma_{ii}^{-1}\right]$$

Therefore,

$$\Sigma^{-1}J(J'\Sigma^{-1}J)^{-1} = \begin{bmatrix}\sigma_{11}^{-1} \\ \sigma_{22}^{-1} \\ \vdots \\ \sigma_{nn}^{-1}\end{bmatrix}\left[\sum_i \sigma_{ii}^{-1}\right]^{-1}$$

This is a matrix expression of the well-known rule that independent observations should be weighted according to the reciprocals of their variances.

[4] Ibbotson [5], chapter 9, pp. 147-156.

[5] Appendix A shows how this solution was obtained.

[6] Butsic [1], pp. 345-347.

[7] Cf. Judge [3], pp. 951-953 on eigenvalues and eigenvectors. Also pp. 960f. Using the SAS IML routine EIGEN, the author decomposed Σ as V'ΛV, where **V** is the eigenvector matrix and Λ is the eigenvalue matrix:

$$\mathbf{V} = \begin{bmatrix} \mathbf{0.3200} & \mathbf{-0.0104} & \mathbf{0.6346} & \mathbf{-0.7034} & \mathbf{0} \\ \mathbf{0.7992} & \mathbf{0.0670} & \mathbf{0.2156} & \mathbf{0.5570} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} \\ \mathbf{-0.0978} & \mathbf{0.9933} & \mathbf{0.0621} & \mathbf{-0.0032} & \mathbf{0} \\ \mathbf{-0.4993} & \mathbf{-0.0940} & \mathbf{0.7396} & \mathbf{0.4414} & \mathbf{0} \end{bmatrix}$$

$$\Lambda = \begin{bmatrix} \mathbf{0.2684} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0.1145} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0.0100} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0.0024} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{-0.0428} \end{bmatrix}$$

V is orthogonal, meaning that **VV'** = **V'V** = **I** (the identity matrix). The problem of optimizing **X'ΣX**, subject to **AX** = **B** can be transformed by a rotation of axes specified by **V**: **X'ΣX** = **X'(V'V)Σ(V'V)X** = **(X'V')(VΣV')(VX)** = **(VX)'(VV'ΛVV')(VX)** = **Y'(Λ)Y**.

And $\mathbf{B} = \mathbf{AX} = \mathbf{A(V'V)X} = \mathbf{(AV')(VX)} = \mathbf{(AV')Y} = \mathbf{CY}$. So the transformed problem is to optimize:

$$\mathbf{Y'\Lambda Y} = \begin{bmatrix} y_1 & y_2 & y_3 & y_4 & y_5 \end{bmatrix} \begin{bmatrix} \mathbf{0.2684} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0.1145} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0.0100} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0.0024} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{-0.0428} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix}$$

subject to $\mathbf{CY} = \mathbf{(AV')Y} = \mathbf{B}$. Let us consider the first constraint, viz., that the surplus could be apportioned in any manner so long as it totaled \$400:

$$\mathbf{CY} = \begin{bmatrix} \mathbf{0.2407} & \mathbf{1.6388} & \mathbf{1} & \mathbf{0.9544} & \mathbf{0.5878} \end{bmatrix} \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \mathbf{y}_3 \\ \mathbf{y}_4 \\ \mathbf{y}_5 \end{bmatrix} = \begin{bmatrix} \mathbf{400} \end{bmatrix}$$

One very obvious choice for **Y** is for the first four elements to be zero, and for the fifth element to be 680.49. At this point $RBC^2 = \mathbf{X'\Sigma X} = \mathbf{Y'\Lambda Y} = [-19{,}799.92]$. The corresponding value for **X** is:

$$\mathbf{X} = \mathbf{V'Y} = \begin{bmatrix} \mathbf{0.3200} & \mathbf{0.7992} & \mathbf{0} & \mathbf{-0.0978} & \mathbf{-0.4993} \\ \mathbf{-0.0104} & \mathbf{0.06696} & \mathbf{0} & \mathbf{0.9933} & \mathbf{-0.0940} \\ \mathbf{0.6346} & \mathbf{0.2156} & \mathbf{0} & \mathbf{0.0621} & \mathbf{0.7396} \\ \mathbf{-0.7034} & \mathbf{0.5570} & \mathbf{0} & \mathbf{-0.0032} & \mathbf{0.4414} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{680.49} \end{bmatrix} = \begin{bmatrix} \mathbf{-339.75} \\ \mathbf{-63.94} \\ \mathbf{503.30} \\ \mathbf{300.39} \\ \mathbf{0} \end{bmatrix}$$

In this situation the company has \$503.30 in affiliates, has borrowed \$63.94 in bonds, and is sold short \$339.75 in stock. So far this is a plausible balance sheet. But the company would have to show (\$300.39) in loss reserve, which is not realistic (sooner or later it may be possible with exotic reinsurance treaties or with insurance futures). However, the point is made that risk-based capital should use a positive definite matrix of correlations.

[8] After the body of this article was written the NAIC published the risk-based capital formula applicable to 12/31/94 balance sheets. Briefly, the NAIC simplified the covariance problem by assuming that off-diagonal covariances are zero. The author cannot guess how true-to-reality this simplification remains; however, it does insure that the variance matrix Σ is positive definite. Cf. NAIC Proceedings 1993 1st Quarter, p. 163.

REFERENCES

1. Butsic, Robert P., "Solvency Measurement for Property-Liability Risk-Based Capital Applications," Insurer Financial Solvency, Casualty Actuarial Society Discussion Paper Program, 1992, p. 311.

2. Eves, Howard, Elementary Matrix Theory (New York: Dover, 1966).

3. Judge, George C., *et al.*, Introduction to the Theory and Practice of Econometrics, 2nd ed (New York: John Wiley & Sons, 1988).

4. Markowitz, Harry W., "Mean-Variance Analysis," The New Palgrave Finance (New York: W. W. Norton, 1989), p. 194.

5. Stocks, Bonds, Bills, and Inflation 1994 Yearbook (Chicago: Ibbotson Associates, 1994).

APPENDIX A

OPTIMIZATION OF A QUADRATIC FORM UNDER CONSTRAINTS

In section 4 was presented the solution to the problem of finding the critical point of a quadratic form subject to constraints. In this appendix the solution will be derived. The derivation requires the use of what is known as a Lagrange multiplier.

The problem is to find the critical point of $\mathbf{X}'\Sigma\mathbf{X}$ subject to $\mathbf{AX} = \mathbf{B}$; where $\mathbf{X}$ is (n x 1), Σ is symmetric (n x n), $\mathbf{A}$ is (m x n), and $\mathbf{B}$ is (m x 1). Let λ be the (m x 1) Lagrangian vector. We form the Lagrangian $\Lambda(\mathbf{X},\lambda) = \mathbf{X}'\Sigma\mathbf{X} + 2\lambda'(\mathbf{AX} - \mathbf{B})$. According to the rules of matrix differentiation (cf. Judge [3], pp. 967-969):

$$\frac{\partial \Lambda}{\partial \mathbf{X}} = 2\Sigma\mathbf{X} + 2\mathbf{A}'\lambda$$

$$\frac{\partial \Lambda}{\partial \lambda} = 2(\mathbf{AX} - \mathbf{B})$$

The critical values $\mathbf{X}_0$ and λ_0 are obtained by setting the derivatives to $\mathbf{0}$. Therefore:

$$\begin{aligned} \Sigma\mathbf{X}_0 + \mathbf{A}'\lambda_0 &= \mathbf{0} \\ \mathbf{AX}_0 \qquad &= \mathbf{B} \end{aligned}$$

where $\mathbf{0}$ in the first equation is an (n x 1) vector. We can combine these two equations into one partitioned-matrix equation:

$$\begin{bmatrix} \Sigma & \mathbf{A}' \\ \mathbf{A} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{X}_0 \\ \lambda_0 \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{B} \end{bmatrix}.$$

If the inverse exists, then the solution is:

$$\begin{bmatrix} \mathbf{X}_0 \\ \lambda_0 \end{bmatrix} = \begin{bmatrix} \Sigma & \mathbf{A}' \\ \mathbf{A} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{0} \\ \mathbf{B} \end{bmatrix}.$$

If Σ is positive definite and rank(**A**) = m, then $\mathbf{A}\Sigma^{-1}\mathbf{A}'$ is a positive definite (m x m) matrix. Therefore, $\mathbf{H} = -(\mathbf{A}\Sigma^{-1}\mathbf{A}')^{-1}$ exists, and $\begin{bmatrix} \Sigma & \mathbf{A}' \\ \mathbf{A} & \mathbf{0} \end{bmatrix}^{-1} = \begin{bmatrix} \Sigma^{-1}(\mathbf{I}_n + \mathbf{A}'\mathbf{H}\mathbf{A}\Sigma^{-1}) & -\Sigma^{-1}\mathbf{A}'\mathbf{H} \\ -\mathbf{H}\mathbf{A}\Sigma^{-1} & \mathbf{H} \end{bmatrix}$. One can verify this last equation by multiplying both matrices, using the definition of **H**, and arriving at $\mathbf{I}_{n+m}$, the identity matrix of dimension n+m. Therefore,

$$\begin{bmatrix} \mathbf{X}_0 \\ \lambda_0 \end{bmatrix} = \begin{bmatrix} \Sigma^{-1}(\mathbf{I}_n + \mathbf{A}'\mathbf{H}\mathbf{A}\Sigma^{-1}) & -\Sigma^{-1}\mathbf{A}'\mathbf{H} \\ -\mathbf{H}\mathbf{A}\Sigma^{-1} & \mathbf{H} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \mathbf{B} \end{bmatrix}, \text{ and}$$

$\mathbf{X}_0 = -\Sigma^{-1}\mathbf{A}'\mathbf{H}\mathbf{B} = \Sigma^{-1}\mathbf{A}'(\mathbf{A}\Sigma^{-1}\mathbf{A}')^{-1}\mathbf{B}$. As shown in section 4, $\mathbf{X}_0$ will be a minimum if Σ is positive definite.