On Mean-Variance Analysis [1]

Yang Li
Rotman School of Management
University of Toronto
105 St. George Street
Toronto, ON, M5S 3E6
Yang.Li10@rotman.utoronto.ca

Traian A. Pirvu
Dept of Mathematics & Statistics
McMaster University
1280 Main Street West
Hamilton, ON, L8S 4K1
tpirvu@math.mcmaster.ca

Abstract

This paper considers the mean variance portfolio management problem. We examine portfolios which contain both primary and derivative securities. The challenge in this context is the well posedness of the optimization problem. We find examples in which this problem is well posed. Numerical experiments provide the efficient frontier. The methodology developed in this paper can be also applied to pricing and hedging in incomplete markets.

JEL classification: C61; G11

Keywords: Mean-Variance, Portfolios of Options, Quadratic Programming.

1 Introduction

The main objective in portfolio management is the tradeoff between risk and return. Markovitz, [6] and [7] studied the problem of maximizing portfolio expected return for a given level of risk, or equivalently minimize risk for a given amount of expected return. One limitation of Markovitz's model is that it considers portfolios of primary assets only.

Recent works consider the optimal management of portfolios containing primary and derivative assets. Here we mention [8] and [1]. [8] introduces a technique of optimizing CVaR (conditional value at risk) of a portfolio. [1] notices that the problem of minimizing CVaR for a portfolio of derivative securities is ill-posed. [1] shows that this predicament can be overcome by including transaction costs.

There are some works which consider portfolio optimization with non-standard asset classes; we recall [2], [3], and [5]. In a continuous time model [2] looks at the problem of maximizing expected exponential utility of terminal wealth, by trading a static position in derivative securities and a dynamic position in stocks. In a one period model [3] analyses the optimal investment and equilibrium pricing of primary and derivative instruments. [5] shows how to approximate a dynamic position in options by a static one and this is done by minimizing the mean-squared error.

By the best of our knowledge this paper is the first work to consider the mean variance Markovitz portfolio management problem in one period model with derivative assets.

[1] Work supported by NSERC grants 371653-09 and MITACS grants 5-26761

We work in a multivariate normally distributed returns framework. The first difficulty encountered in such problems stems from the nonlinearity of derivative prices. This can be overcome by considering delta gamma or related approximations. It is well known in the industry practice that this approximation performs well for small time intervals. By performing the delta gamma approximation the portfolio management problem is reduced to a quadratic program. The second difficulty is the lack of convexity for the quadratic program. We find two examples in which the quadratic program is convex: a portfolio of instruments in which every instrument is written on one underlying only; a second example is a portfolio which contains only two instruments.

The results of our paper can be applied to the problem of pricing and hedging in incomplete markets. For instance we can consider instruments written on nontradable factors (e.g. temperature) and they can be hedged with tradable instruments which are highly correlated (this procedure is called cross hedging). Take as an example weather derivatives (e.g. HDD or CDD); energy prices are considered as the traded correlated instrument (in California a highly correlation can be observed between temperature and energy prices). Perfect hedging is not possible in this paradigm. Minimizing the variance of the hedging error can be captured as a special case of mean variance optimization problem for a portfolio of primary and derivative instruments. A survey paper on mean-variance hedging and mean-variance portfolio selection is [9].

Another possible application of our results is the hedging of long maturity instruments with short maturities ones. As it is well known, the market for long maturity instruments is illiquid, thus the issuers use (static) hedging portfolios of the more liquid short maturity instruments. The interested reader can find out more about this in [4].

The paper is organized as follows: Section 2 presents the model. Section 3 introduces the delta gamma approximation. Section 4 presents the reduction to quadratic programs. Section 5 shows examples in which the quadratic programs are convex. Section 6 presents the numerical results. Section 7 is an application to pricing and hedging in incomplete markets. The paper ends with an appendix.

2 The Model

Portfolios returns are derived from the return of individual positions. In practice, it is not good to model the positions individually because of their correlations. If we have m instruments in our portfolio we would need m separate volatilities plus data on $\frac{m(m-1)}{2}$ correlations, so in total $\frac{m(m+1)}{2}$ pieces of information. This is hard to get for large m.

The resolution is to map our m instruments onto a smaller number of n risk factors. The mapping can be nonlinear (e.g. BS for option). Let us assume that the factors are represented by a stochastic vector process $S = (S_1, S_2, \cdots, S_n)$, which at all times $t \in (0, \infty)$ is assumed to be of the form

$$S_t = \mu t + \Sigma W_t. \tag{2.1}$$

Here μ is the vector of returns, Σ is the variance-covariance matrix, and W_t is a standard Brownian motion on a canonical probability space $(\Omega, \mathcal{F}_t, \mathcal{F})$. The value of portfolio at time t, denoted $V(S,t)$, is of the form

$$V(S,t) = \sum_{k=1}^{m} x_k(t) V_k(S,t), \tag{2.2}$$

where $V_k(S,t),\ k=1,\cdots m$, represents the value of the individual instruments (mapped onto the risk factors), and $x_k(t),\ k=1,\cdots m$ stands for the number of shares of instrument k held in the portfolio at time t. At time t, we choose the portfolio mix $x_k(t),\ k=1,\cdots m$ such that the portfolio return ΔV over time interval $[t, t+\Delta t]$

$$\Delta V = V(S+\Delta s, t+\Delta t) - V(S,t), \tag{2.3}$$

is optimized in a way which is described below. It turns out to be more convenient to work with the vector of actual proportions of wealth invested in the different assets. Thus, at time $t \in (0,\infty)$, we introduce the portfolio weights $w_k(t),\ k=1,\cdots m$, by

$$w_k(t) = \frac{x_k(t)}{V(S,t)},\ k=1,\cdots m. \tag{2.4}$$

In the following we posit the following Markowitz mean-variance type problem; given some exogenous benchmark return $r_e(t)$, at time t an investor wants to choose among all portfolios having the same return $r_e(t)$, the one that has the minimal variance $\mathrm{Var}(\Delta V)$:

$$\begin{array}{lcl} \text{(P1)} & \displaystyle\min_w & \mathrm{Var}(\Delta V) \\ & \text{such that} & \mathrm{E}(\Delta V) = r_e(t), \\ & & \displaystyle\sum_{k=1}^m w_k(t) V_k(S,t) = 1. \end{array}$$

Another possible portfolio management problem is to choose the portfolio with the minimal variance:

$$\begin{array}{lcl} \text{(P2)} & \displaystyle\min_w & \mathrm{Var}(\Delta V) \\ & & \displaystyle\sum_{k=1}^m w_k(t) V_k(S,t) = 1, \end{array}$$

There are some difficulties in solving (P1) and (P2). First, we might be short of information of the moments of ΔV. Because ΔV nonlinearly depends on the change of asset prices, it is not obvious what distribution ΔV would follow even if we perfectly learn the p.d.f of ΔS. The situation would not get much better if we only require the moment information of ΔV. The integration for moments might be still hard to calculate explicitly. One way what of this predicament is to use delta gamma approximation.

3 Delta-Gamma Approximation

This is a second order Taylor expansion

$$\Delta V \approx\ \delta V = \frac{\partial V}{\partial t}\Delta t + \delta^T \Delta S + \frac{1}{2}\Delta S^T \Gamma \Delta S, \tag{3.1}$$

where

$$\delta_i = \frac{\partial V}{\partial S_i}, \qquad \Gamma_{ij} = \frac{\partial^2 V}{\partial S_i \partial S_j}, \quad i = 1,\cdots n.$$

Since

$$V(S,t) = \sum_{k=1}^m x_k(t) V_k(S,t),$$

then

$$\delta_i = \frac{\partial V}{\partial S_i} = \sum_{k=1}^{m} x_k(t)\delta_i^k,\ \ \delta_i^k := \frac{\partial V_k}{\partial S_i},\ \ i = 1, \cdots n,\ \ k = 1, \cdots m, \tag{3.2}$$

$$\Gamma_{ij} = \frac{\partial^2 V}{\partial S_i \partial S_j} = \sum_{k=1}^{m} x_k(t)\Gamma_{ij}^k,\ \Gamma_{ij}^k := \frac{\partial^2 V_k}{\partial S_i \partial S_j},\ i = 1, \cdots n,\ j = 1, \cdots n,\ k = 1, \cdots m. \tag{3.3}$$

It is well known that this approximation performs well as long as the time interval Δt is not too big. At this point we formulate the approximated versions of (P1) and (P2), as follows:

$$\begin{array}{llll} \text{(P3)} & \min\limits_{w} & \mathrm{Var}(\delta V) \\ & \text{such that} & \mathrm{E}(\delta V) = r_e(t), \\ & & \sum\limits_{k=1}^{m} w_k(t)V_k(S,t) = 1, \end{array}$$

$$\begin{array}{llll} \text{(P4)} & \min\limits_{w} & \mathrm{Var}(\delta V) \\ & & \sum\limits_{k=1}^{m} w_k(t)V_k(S,t) = 1. \end{array}$$

The next step is to reduce (P3) and (P4) to quadratic programs and this is done in the next section.

4 Quadratic Programs

Let us first consider the case of one asset, $m = 1$. In the light of (2.1), $\Delta S \sim \mathcal{N}(\mu, \Sigma\sqrt{\Delta t})$. For computational convenience we assume μ is the zero vector and $\Delta t = 1$. Next, replace the vector of correlated normals, ΔS, with the vector of independent normals $Z \sim \mathcal{N}(0, I)$. This is done by setting

$$\Delta S = CZ \qquad \text{with} \qquad CC^T = \Sigma.$$

In terms of Z, the quadratic approximation of ΔV becomes

$$\Delta V \approx \delta V = a + (C^T\delta)^T Z + \frac{1}{2} Z^T (C^T \Gamma C) Z,$$

with

$$a = \frac{\partial V}{\partial t}\Delta t. \tag{4.1}$$

At this point it is convenient to choose the matrix C to diagonalize the quadratic term in the above expression and this is done as follows. Let $\tilde{C}$ be a square matrix such that

$$\tilde{C}\tilde{C}^T = \Sigma \tag{4.2}$$

(e.g., the one given by the Cholesky factorization). The matrix $\frac{1}{2}\tilde{C}^T\Gamma\tilde{C}$ is symmetric and thus admits the representation (more about this can be found in the appendix)

$$\frac{1}{2}\tilde{C}^T\Gamma\tilde{C} = U\Lambda U^T, \tag{4.3}$$

where $\Lambda = \text{diag}(\lambda_1, \cdots, \lambda_n)$, and U is an orthogonal matrix such that $UU^T = I$. Next, set $C = \tilde{C}U$ and observe that

$$CC^T = \tilde{C}UU^T\tilde{C}^T = \Sigma, \tag{4.4}$$

$$\frac{1}{2}C^T\Gamma C = \frac{1}{2}U^T(\tilde{C}^T\Gamma\tilde{C})U = U^T(U\Lambda U^T)U = \Lambda.$$

Thus, with

$$b = C^T\delta, \tag{4.5}$$

we get

$$\Delta V \approx \delta V = a + b^T Z + Z^T \Lambda Z := Y.$$

4.1 Moment Generating Function

In this subsection, we explore the moment generating function of Y and further derive the mean and variance of Y. Since the random variable Y is student distributed, it is well known that

$$\mathrm{E}(\theta Y) = \exp(\eta(\theta)), \tag{4.6}$$

where

$$\eta(\theta) = a\theta + \sum_{j=1}^{n} \eta_j(\theta) = a\theta + \sum_{j=1}^{n} \frac{1}{2}\left(\frac{\theta^2 b_j^2}{1-2\theta\lambda_j} - \log(1-2\theta\lambda_j)\right), \tag{4.7}$$

for all θ satisfying $\max_j \theta\lambda_j < \frac{1}{2}$. Direct computations lead to

$$\begin{aligned}
\frac{d\left(e^{\eta(\theta)}\right)}{d\theta} &= \exp(\eta(\theta))\frac{d\eta}{d\theta} \\
&= \exp(\eta(\theta))\left[a + \frac{1}{2}\sum_{j=1}^{n}\left(\frac{2\theta b_j^2(1-2\theta\lambda_j) - \theta^2(-2\lambda_j)}{(1-2\theta\lambda_j)^2} - \frac{-2\lambda_j}{1-2\theta\lambda_j}\right)\right]
\end{aligned}$$

and

$$\begin{aligned}
\frac{d^2\left(e^{\eta(\theta)}\right)}{d\theta^2} &= \exp(\eta(\theta))\left[a + \frac{1}{2}\sum_{j=1}^{n}\left(\frac{2\theta b_j^2(1-2\theta\lambda_j) - \theta^2(-2\lambda_j)}{(1-2\theta\lambda_j)^2} - \frac{-2\lambda_j}{1-2\theta\lambda_j}\right)\right]^2 \\
&\quad + \exp(\eta(\theta))\left[\tfrac{1}{2}\sum_{j=1}^{n}\left(2\lambda_j\frac{-2\lambda_j}{-(1-2\theta\lambda_j)^2} + \right.\right. \\
&\left.\left.\frac{(2b_j^2 - 8\theta b_j^2\lambda_j + 4\theta b_j^2\lambda_j)(1-2\theta\lambda_j)^2 - (2\theta b_j^2(1-2\theta\lambda_j) + \theta^2 b_j^2 2\lambda_j)((-2\lambda_j)2(1-2\theta\lambda_j))}{(1-2\theta\lambda_j)^4}\right)\right]
\end{aligned}$$

Thus, the first and second moments of Y are

$$\mathrm{E}(Y) = \left.\frac{d\left(e^{\eta(\theta)}\right)}{d\theta}\right|_{\theta=0} = a + \sum_{j=1}^{n}\lambda_j$$

and

$$\mathrm{E}(Y^2) = \left.\frac{d^2\left(e^{\eta(\theta)}\right)}{d\theta^2}\right|_{\theta=0} = (a + \sum_{j=1}^{n}\lambda_j)^2 + \sum_{j=1}^{n}(b_j^2 + 2\lambda_j^2).$$

Hence,

$$\mathrm{Var}(Y) = \mathrm{E}(Y^2) - \mathrm{E}^2(Y) = \sum_{j=1}^{n}(b_j^2 + 2\lambda_j^2).$$

In order to ease the notations we assume that $V(S,t) = 1$, so the vector of shares x equals the vector of proportions w (also notice that for simplicity we dropped the t dependence of w). Next, we would like to express the mean and variance of Y in terms of x. In the light of (4.3), (4.2), (3.3) and (8.1) it follows that

$$\begin{aligned}
\mathrm{E}(Y) = a + \sum_{j=1}^{n}\lambda_j &= a + \mathrm{tr}\left(U\Lambda U^T\right) \\
&= a + \frac{1}{2}\mathrm{tr}\left(\tilde{C}^T\Gamma\tilde{C}\right) \\
&= a + \frac{1}{2}\mathrm{tr}\left(\sum_{j=1}^{n} x_j\Gamma_j\Sigma\right) \\
&= a + x^T p,
\end{aligned} \tag{4.8}$$

where the vector p is defined by

$$p := \frac{1}{2}\left(\mathrm{tr}\left(\Gamma^1\Sigma\right), \mathrm{tr}\left(\Gamma^2\Sigma\right), \ldots, \mathrm{tr}\left(\Gamma^n\Sigma\right)\right)^T.$$

As for the variance, recall that with b of (4.5) it follows that (see (3.2) and (4.4))

$$\sum_{k=1}^{m} b_k^2 = b^T b = (C^T\delta)^T C^T\delta = \delta^T C^T C\delta = \frac{1}{2}x^T\hat{\Sigma}x, \tag{4.9}$$

where

$$\hat{\Sigma} = 2(\delta^1, \cdots, \delta^n)\ \Sigma\ (\delta^1, \cdots, \delta^n)^T.$$

$\hat{\Sigma}$ is positive semidefinite by Lemma 8.1 (see appendix). Next, in the light of (4.4) and (8.1) it follows that

$$\begin{aligned}
\sum_{k=1}^{m}\lambda_j^2 &= \frac{1}{4}\mathrm{tr}\left((C^T\Gamma C)^T(C^T\Gamma C)\right) \\
&= \frac{1}{4}\mathrm{tr}\left(\Gamma CC^T\Gamma CC^T\right) \\
&= \frac{1}{4}\mathrm{tr}\left(\Gamma\Sigma\Gamma\Sigma\right) \\
&= \frac{1}{4}\mathrm{tr}\left(\left(\sum_{j=1}^{n} x_j\Gamma^j\Sigma\right)^2\right) \\
&= \frac{1}{4}\left(\sum_{j=1}^{n} x_j^2\mathrm{tr}\left((\Gamma^j)^2\Sigma^2\right) + 2\sum_{i\neq j} x_i x_j \mathrm{tr}\left(\Gamma^i\Sigma\Gamma^j\Sigma\right)\right) \\
&= \frac{1}{4}x^T Q x,
\end{aligned}$$

where the matrix Q is defined by

$$Q_{ij} = \mathrm{tr}\left(\Gamma^i \Sigma \Gamma^j \Sigma\right),\ i = 1, \cdots n,\ j = 1, \cdots n. \tag{4.10}$$

Therefore, we end up with

$$\mathrm{Var}(Y) = \sum_{j=k}^{m} (b_k^2 + 2\lambda_k^2) = \frac{1}{2} x^T (\hat{\Sigma} + Q) x. \tag{4.11}$$

Thus, from (4.8) and (4.11), the portfolio problem (P3) (recall that $x = w$) becomes

$$\begin{array}{lll} \text{(P5)} & \min\limits_{x} & \frac{1}{2} x^T (\hat{\Sigma} + Q) x \\ & \text{s.t.} & a + x^T p = r_e, \\ & & \sum\limits_{k=1}^{n} V_k(t, S) x_k = 1. \end{array}$$

and (P4) turns into

$$\begin{array}{lll} \text{(P6)} & \min\limits_{x} & \frac{1}{2} x^T (\hat{\Sigma} + Q) x \\ & \text{s.t.} & \sum\limits_{k=1}^{n} V_k(t, S) x_k = 1. \end{array}$$

It turns out that the problem (P5) has a similar form with the classical mean variance portfolio problem. However, the essential difference is that (P5) is not always tractable. Depending on the data, (P5) could be either a convex quadratic programming or NP-hard. The following is a well known result in quadratic programming.

Theorem 4.1. *(P5) is a convex quadratic program (thus solvable in polynomial time) as long as the matrix $\hat{\Sigma} + Q$ is positive semidefinite. Else, (P5) is NP-hard.*

In the following we want to find examples in which (P5) is convex, and this is done in the next section.

5 Convex Quadratic Programs

As Theorem 4.1 states, (P5) is not always tractable. We found two instances in which (P5) is convex.

5.1 Case 1

Let us consider a portfolio in which every instrument is a map of one factor (possible different from instrument to instrument) only. Take for example a portfolio of options written on one (possible different) underlying. Thus,

$$V(S, t) = \sum_{k=1}^{m} x_k V_k(S_k, t).$$

Since V_k only depends on the factor S_k, the matrix Γ^k (see (3.3)) has a nonzero element (Γ^k_{kk}) only. We prove that in this case (P5) is a convex quadratic program. Given this special structure it follows that Q_{ij} of (4.10) becomes

$$Q_{ij} = \mathrm{tr}\left(\Gamma^i \Sigma \Gamma^j \Sigma\right) = \Gamma^i_{ii}\Gamma^j_{jj}\Sigma_{ij}\Sigma_{ji}. \tag{5.1}$$

After some algebra manipulation one gets

$$Q = D^T(\Sigma \circ \Sigma)D, \tag{5.2}$$

where $D = \mathrm{Diag}(\Gamma_1, \Gamma_2, \ldots, \Gamma_m)$ and $\circ$ denotes the Hadamard product. Recall that for two matrices $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{m\times n}$, the Hadamard product is the matrix $A \circ B$ with the entries

$$(A \circ B)_{ij} := A_{ij}B_{ij}.$$

According to Schur Product Theorem (see appendix) it follows that $\Sigma \circ \Sigma$ is positive semidefinite. Next, by Lemma 8.1 it follows that $D^T(\Sigma \circ \Sigma)D$ is also positive semidefinite. At this point we can state the result

Theorem 5.1. *(P5) is a convex quadratic program.*

Proof. So far we proved that the matrices $\bar{\Sigma}$ and Q are positive semidefinite. Hence (P5) is a convex quadratic program, since $\bar{\Sigma} + Q$ is positive semidefinite. $\square$

Let us move to the second case in which we can establish convexity of (P5).

Case 2

Let us consider a portfolio containing two instruments only. That is,

$$V(S,t) = \sum_{k=1}^{2} x_k V_k(S,t).$$

Thus

$$\begin{aligned} Q &= \begin{pmatrix} \mathrm{tr}\left(\Gamma^1\Sigma\Gamma^1\Sigma\right) & \mathrm{tr}\left(\Gamma^1\Sigma\Gamma^2\Sigma\right) \\ \mathrm{tr}\left(\Gamma^2\Sigma\Gamma^1\Sigma\right) & \mathrm{tr}\left(\Gamma^2\Sigma\Gamma^2\Sigma\right) \end{pmatrix} \\ &= \begin{pmatrix} \mathrm{tr}\left((C^T\Gamma^1C)^TC^T\Gamma^1C\right) & \mathrm{tr}\left((C^T\Gamma^1C)^TC^T\Gamma^2C\right) \\ \mathrm{tr}\left((C^T\Gamma^1C)^TC^T\Gamma^2C\right) & \mathrm{tr}\left((C^T\Gamma^2C)^TC^T\Gamma^2C\right) \end{pmatrix}. \end{aligned}$$

We claim that the matrix Q is positive semidefinite. First, $\mathrm{tr}\left((C^T\Gamma_1C)^TC^T\Gamma_1C\right)$ is nonnegative. For the determinant condition we apply the Cauchy-Schwarz inequality

$$|\langle X_1, X_2\rangle| \le \|X_1\| \cdot \|X_2\|,$$

with $X_1 = C^T\Gamma_1C$ and $X_2 = C^T\Gamma_2C$. Here the inner product $\langle X_1, X_2\rangle$ is defined by

$$\langle X_1, X_2\rangle = \mathrm{tr}\left(X_1X_2^T\right).$$

Consequently, $\hat{\Sigma} + Q$ is positive definite, which lead to the following Theorem.

Theorem 5.2. *(P5) is a convex quadratic program.*

6 Numerical Results

Assume that there is one riskless asset, SC 91 Day T-Bills, and four risky assets, SC Universe Bonds, S&P/TSX, S&P 500, and MSCI EAFE. Further, assume that the arithmetic returns are multivariate normally distributed. The expected return, standard deviation and the covariance matrix is obtained from the historical data (this is taken from Jan. 1993 to Dec. 2002).[2] Next, consider a vanilla call option on SC Universe Bonds, a vanilla call option on S&P/TSX, a binary put option on BiS&P 500, and a binary put option on MSCI EAFE. The initial asset prices, strike prices and maturity dates are given in Table 3. The derivatives are priced using BS type formulas.

Table 1: Expected Return and Standard Deviation

	SC91TBIL	SCUNOVER	SP/TSX	SP500	MSEAFEC
Return	4.85%	8.91%	9.07%	11.68%	6.54%
STDEV	0.44%	5.08%	16.80%	13.90%	14.35%

Table 2: Covariance Matrix

	SC91TBIL	SCUNOVER	SP/TSX	SP500	MSEAFEC
SC91TBIL	0.00%	0.01%	0.01%	0.01%	0.00%
SCUNOVER	0.01%	0.26%	0.24%	0.10%	0.02%
SP/TSX	0.01%	0.24%	2.82%	1.64%	1.49%
SP500	0.01%	0.10%	1.64%	1.93%	1.30%
MSEAFEC	0.00%	0.02%	1.49%	1.30%	2.06%

Table 3: Initial prices, strike prices, and expire date

	SCUNOVER	SP/TSX	SP500	MSEAFEC
Initial Price	100	50	80	100
Strike Price	80	51.25	100	150
Expire Date (days)	20	40	60	80

7 Quadratic Hedging

The results we established so far can also be applied to hedging. The motivation comes from incomplete markets. Indeed, financial markets are fundamentally incomplete. It is well known that in incomplete markets perfect hedging is not possible. One way to solve this problem is to consider quadratic hedging; that is, minimize the variance of the hedging error. Let F be a payoff of the form $F = V_1(S_{t+\Delta t}, t + \Delta t)$, for some map V_1. We would like to hedge this payoff by some instruments which are of the form $V_k(S,t)$, $k = 2, \cdots, l$ (with l possible less than n, whence the incompleteness). For simplicity assume that in this market borrowing and lending of cash is done

[2] Data from http://www.math.mcmaster.ca/~grasselli/john.pdf. For the covariance matrix, we only take 4 decimals.

at zero interest rate (this can be easily achieved if one takes the zero coupon bonds as numeraire). Given the number of shares $(x_1, x_2, \cdots, x_l)$ in the hedging portfolio, the hedging error is

$$-\sum_{k=1}^{l+1} x_k \Delta V_k(S,t),$$

with $x_1 = -1$, and $\Delta V_{l+1}(S,t) = 1$. Therefore, the problem of minimizing the variance of hedging error is of the form (P2). The initial amount needed to finance the hedging portfolio is

$$x_{l+1} + V_1(S_t, t).$$

8 Appendix

8.1 Trace and its Properties

In linear algebra, the trace of a square matrix A is defined to be the sum of the elements on the main diagonal (the diagonal from the upper left to the lower right) of A, i.e.,

$$\mathrm{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn} = \sum_{i=1}^{n} a_{ii}.$$

Property 8.1. If $A \in \mathcal{R}^{m\times n}$ and $B \in \mathcal{R}^{n\times m}$, then

$$\mathrm{tr}(AB) = \mathrm{tr}(BA). \tag{8.1}$$

Property 8.2. If P is an invertible matrix, then

$$\mathrm{tr}(P^{-1}AP) = \mathrm{tr}((AP)P^{-1}) = \mathrm{tr}(A). \tag{8.2}$$

Property 8.3. If $A \in \mathcal{R}^{n\times n}$ with real or complex entries and if $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the (complex and distinct) eigenvalues of A (listed according to their algebraic multiplicities), then

$$\mathrm{tr}(A) = \sum_{i=1}^{n} \lambda_i. \tag{8.3}$$

8.2 Symmetric Matrices and Semidefinite Positive Matrices

Property 8.4 (Eigenvalue Decomposition). Let M be a symmetric real matrix. Then, there exists an *Eigenvalue Decomposition* such that

$$M = Q\Lambda Q^T = \sum_{i=1}^{n} \lambda_i q_i q_i^T, \tag{8.4}$$

where Λ is a diagonal matrix with all the eigenvalues of M along its diagonal, Q is an orthonormal matrix, i.e., $QQ^T = I$, and each column q_i of Q is an eigenvector of M corresponding to the eigenvalue λ_i.

Lemma 8.1. *Let $P \in R^{n\times n}$ and $A \in R^{n\times n}$ be a semidefinite positive matrix. Then the matrix PAP^T is also semidefinite positive..*

Proof. According to the definition, we have

$$x^T PAP^T x = (P^T x)A(P^T x) \geq 0,$$

for any $x \neq 0$. $\square$

8.3 Schur Product Theorem

Property 8.5 (Schur Product Theorem). Suppose $A, B \in \mathcal{R}^{n \times n}$ are positive semidenite matrices. Then $A \circ B$ is also positive semidenite.

□

References

[1] Alexander, S., Coleman, T., F., and Li, Y. (2006), Minimizing CVaR and VaR for a portfolio of derivatives, *Journal of Banking & Finance*, **30**, 583-605.

[2] Aytac, I., Jonsson, M., and Sircar, R. (2005), Optimal Investment with Derivative Securities *Finance and Stochastics*, **9**, 585-595.

[3] Carr, P. and Madan, D. (2001) Optimal positioning in derivative securities, *Quantitative Finance*, **1**, 19-37.

[4] Carr, P. and Wu, L. (2004) Static Hedging of Standard Options, *Preprint*.

[5] Haugh, M. B., and Lo, A., W. (2001), Asset allocation and derivatives *Quantitative Finance* , **1**, 45-72.

[6] Markowitz, H. (1952) Portfolio Selection, *The Journal of Finance* **7**, 77-91.

[7] Markowitz, H. (1991) Foundations of Portfolio Theory *Journal of Finance*, **46**, 469-477.

[8] Rockafellar, R., T. and Uryasev, S. (2000) Optimization of conditional value-at-risk, *Journal of Risk*, **2**, 21-41.

[9] Schweizer, M. (2010) Mean-Variance Hedging, *Encyclopedia of Quantitative Finance*, 1177-1181.