Optimization Methods in Finance **(EPFL, Fall 2010)**

Lecture 5: Mean variance portfolio optimization & Lagrange Duality

20.10.2010

Lecturer: Prof. Friedrich Eisenbrand *Scribe: Julien Damond*

Recall

First we recall some notions that we saw at the end of the last lecture. We consider n assets and denote the *expected rate of return* of asset $i = 1, \ldots, n$ as $\overline{r_i}$. The expected rate of return of portfolio $x \in \Sigma^n$ is given by

$$\overline{r} = \sum_{i=1}^{n} x_i \overline{r_i}.$$

The variance of the portfolio return is defined as

$$\begin{aligned}
\mathbf{Var}\left[\sum_{i=1}^n x_i r_i\right] &= \sum_{i,j} \mathbf{Cov}(x_i r_i, x_j r_j) && = \sum_{i,j} \left(\overline{x_i r_i x_j r_j} - \overline{x_i r_i}\ \overline{x_j r_j}\right) \\
&= \sum_{i,j} \left(x_i x_j \overline{r_i r_j} - x_i \overline{r_i} x_j \overline{r_j}\right) && = \sum_{i,j} x_i \mathbf{Cov}(r_i, r_j) x_j \\
&= x^T Q x
\end{aligned}$$

where $Q(i,j) = \mathbf{Cov}(r_i, r_j)$ is called *covariance matrix*.

Basic version of Markovitz portfolio optimization problem

The problem of computing a minimum variance portfolio among those portfolios with expected return bigger than a given r can be formulated as follows

$$\min\ x^T Q x \tag{1}$$

$$\sum_{i=1}^{n} x_i \overline{r_i} \geq r \tag{2}$$

$$x \in \Sigma^n, \tag{3}$$

where r is the required minimum return. Problem (1) is called the basic version of the Markovitz portfolio optimization problem.

We present two properties of the covariance matrix.

Properties of Q

- $Q^T = Q$ (Symmetry)
- $x^T Qx \geq 0\ \forall x \in \mathbb{R}^n$ (Pos. semidefinite)

Definition 1. A symmetric matrix $Q \in \mathbb{R}^n$ is called *positive semidefinite* if

$$\forall x \in \mathbb{R}^n : \ x^T Qx \geq 0.$$

We also write $Q \succeq 0$.

Recall a Theorem from linear algebra:

Theorem 2. *For $A \in \mathbb{R}^{n\times n}$ the following statements are equivalent*

1. $A \succeq 0$
2. *There is a matrix $Q \in \mathbb{R}^{n\times n}$ with $A = Q^T Q$.*

We give a lemma which will be useful to find the minimum of (1).

Lemma 3. *The function $f : \mathbb{R}^n \to \mathbb{R}$ defined by*

$$f(x) = \ x^T Qx$$

with $Q \succeq 0$ is a convex function.

In order to prove this lemma, we present two results. (They will be proved during the exercises).

- If $f(x)$ is convex and non-negative, then $f(x)^2$ is also convex.
- If $f(x)$ is convex, then $f(Ax+b)$ is convex for any $A \in \mathbb{R}^{m\times n},\ b \in \mathbb{R}^m$

Let us return to our lemma. We can write

$$x^T Qx = x^T A^T Ax = (Ax)^2 = ||Ax||^2$$

for a matrix $A \in \mathbb{R}^{n\times n}$. Here $||\cdot||$ denotes the Euclidean norm.

With the two previous statements, it remains to show that $||\cdot||$ is convex. Let $||\cdot||^*$ be any norm. Then we have by definition of a norm

$$||\lambda x||^* = |\lambda| \cdot ||x||^*,$$

the triangle inequality provides

$$||\lambda x + (1-\lambda)y||^* \leq ||\lambda x||^* + ||(1-\lambda)y||^* = \lambda\ ||x||^* + (1-\lambda)||y||^* \ \text{ if } \lambda \in [0,1].$$

This in fact shows that *any* norm is convex and thus proves the lemma. We define a new function that is composed of elements of (1)

$$g(\lambda) = \min_{x \in \Sigma^n} \left\{ x^T Q x + \lambda \left(r - \sum_{i=1}^{n} x_i \bar{r_i} \right) \right\}$$

for $\lambda \geq 0$.

This function returns the minimum of (1) but adds a penalty when the constraint is not satisfied. We will consider the properties of this function in detail. First we have an inequality that is presented in the following lemma.

Lemma 4. *For all $\lambda \geq 0 : \; g(\lambda) \leq p^*$ where p^* is an optimum value of (1).*

Proof. Let $\lambda \geq 0$ and suppose $\bar{x}$ is an optimum solution of (1). We have

$$g(\lambda) \leq \bar{x}^T Q \bar{x} + \overbrace{\lambda \underbrace{\left(r - \sum_{i=1}^{n} \bar{x_i}\, \bar{r_i} \right)}_{\leq 0}}^{\leq 0} \leq \bar{x}^T Q \bar{x} = p^*.$$

□

A question that one can ask is, is there a value of λ for which the inequality becomes an equality. The answer is given by the following lemma

Lemma 5. *There exists a $\lambda^* \geq 0$ such that $g(\lambda^*) = p^*$.*

The proof will be made in a more general setting later. We give a last result about the function g.

Lemma 6. *The function $g(\lambda)$ defined as before is concave.*

Note that a λ^* for which the equality holds can be found by binary search, exploiting the concavity of g.

Convex optimization problems and duality

The convex optimization problem (C.O.P.) is defined as follows

$$\begin{array}{rcl} \min \;\; f_0(x), & & \\ \text{s.t. } f_i(x) & \leq & 0 \quad \forall i = 1, \ldots, m, \\ x & \in & D, \end{array} \tag{4}$$

where $f_0, \ldots, f_m : \mathbb{R}^m \to \mathbb{R}$ are convex functions and D is a convex domain (like Σ^N). Actually the C.O.P. is a more abstract way to express the Markovitz problem. For resolving the C.O.P. we use another way to express the problem, namely the Lagrangian.

Definition 7. We define the *Lagrangian* as follows

$$L(x,\lambda) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) \quad \text{for } \lambda \in \mathbb{R}_+^m,$$

we also define

$$g(\lambda) = \inf_{x \in D} L(x,\lambda).$$

We can now generalize Lemma 4 as follows:

Theorem 8 (Weak duality)**.** *Suppose (4) has a feasible solution with value p^*. Then*

$$\forall \lambda \in \mathbb{R}_+^m : \, g(\lambda) \le p^*.$$

Proof. Let x^* be a feasible solution of (4). Then $x^* \in D$ and

$$g(\lambda) \le L(x^*,\lambda) = f_0(x^*) + \underbrace{\sum_{i=1}^{m} \lambda_i \underbrace{f_i(x^*)}_{\le 0}}_{\le 0} \le f_0(x^*) = p^*.$$

□

Lagrange dual problem

We know we can transform an max problem into a min problem (and vice versa) by writing

$$\max_{\lambda \ge 0} g(\lambda) = -\min_{\lambda \ge 0} -g(\lambda).$$

Let us make an observation. The function $g(\lambda)$ is concave ($\iff$ $-g(\lambda)$ is convex) and thus Lagrange dual problem is convex optimization problem and this holds regardless of the f_i beeing convex.
To see this, let $\theta \in [0,1]$ and $\lambda_1, \lambda_2 \ge \mathbf{0}$, $x \in D$. We have

$$\begin{aligned} L(x, \theta\lambda_1 + (1-\theta)\lambda_2) &= f_0(x) + \theta \textstyle\sum_{i=1}^m \lambda_i^1 f_i(x) + (1-\theta) \sum_{i=1}^m \lambda_i^2 f_i(x) \\ &= \theta L(x,\lambda_1) + (1-\theta) L(x,\lambda_2). \end{aligned}$$

Finally we have

$$\begin{aligned} g(\theta\lambda_1 + (1-\theta)\lambda_2) &= \inf_{x \in D} L(x, \theta\lambda_1 + (1-\theta)\lambda_2) \\ &= \inf_{x \in D} (L(\theta L(x,\lambda_1) + (1-\theta)L(x,\lambda_2)) \\ &\ge \theta \inf_{x \in D} L(x,\lambda_1) + (1-\theta) \inf_{x \in D} L(x,\lambda_2) \\ &= \theta g(\lambda_1) + (1-\theta) g(\lambda_2). \end{aligned}$$

One can ask when is the optimum value of the Lagrange dual equal to the optimum value of C.O.P ? Well, this is true if the following condition is satisfied:

Definition 9 (Slater's Condition)**.** Suppose D and f_i are convex for $i = 1,2,\dots,m$ and suppose that there exists $x^* \in D$ with $f_i(x^*) < 0$. Then (4) satisfies the *Slater's condition*.

For a more general versions, see [1].

Theorem 10. *If (4) satisfies the Slater's condition and (4) has an optimal solution of value p^*, then*

$$\max_{\lambda \geq 0} g(\lambda) \quad = \min\{\, f_0(x) \mid f_i(x) \leq 0\ \forall 1,\dots,m,\ x \in D\} \quad = p^*.$$

In order to prove the theorem we recall another Theorem, the *separation theorem*.

Theorem 11 (Separation theorem)**.** *If $A, B \in \mathbb{R}^n$ are convex sets with $A \cap B = \emptyset$. Then there exists a hyperplane $a^T x = \beta$ ($a \neq \mathbf{0}$) such that*

$$a^T a^* \leq \beta,\ \forall a^* \in A \ \text{ and } a^T b^* \geq \beta\ \forall b^* \in B.$$

The proof of the separation theorem can be found e.g. in [2]. Intuitively, the theorem says that if we have two convex set with an empty intersection, then there exists an hyperplane that separates the space into two parts, each one contains only one of the two sets.

We focus ourself on the proof of Theorem 10.

Proof. Define

$$A = \{(u,t) \in \mathbb{R}^{m+1} : \exists x \in D \text{ with } f_i(x) \leq u_i,\ i = 1,2,\dots,m,\ f_0(x) \leq t\}.$$

Observe that the set A is convex.

Let $\bar{x} \in D$ satisfy $f_i(\bar{x}) < 0,\ i = 1,2,\dots,m$ and define $B = \{(0,s) \in \mathbb{R}^{n+1} :\ s < p^*\}$, the set B is also convex. Moreover $A \cap B = \emptyset$ since if $(0,s) \in A \cap B$ then $\exists x \in D$ with $f_i(x) \leq 0,\ i = 1,2,\dots,m,\ f_0(x) \leq s < p^*$ this is a contradiction to p^* beeing optimum value of convex minimization problem.

The separation theorem provides that $\exists (\lambda,\mu) \neq \mathbf{0}$ and $\alpha \in \mathbb{R}$ such that

$$(u,t) \in A \implies \lambda^T u + \mu t \geq \alpha, \tag{5}$$

$$(u,t) \in B \implies \lambda^T u + \mu t \leq \alpha. \tag{6}$$

See Figure 1 for an illustration. We claim that the components of λ and μ have to be bigger or equal to 0. Suppose not. Note that $(u,t) \in A \implies (u,t) + \vartheta e_i \in A$ where $\vartheta \in \mathbb{R}_+$ and e_i is the i-th unit vector. If $\lambda_i < 0$ then $(\lambda,\mu)^T \big((\mu,t) + \vartheta e_i\big) \xrightarrow[\vartheta \to \infty]{} -\infty$. But this is a contradiction to $(\lambda,u)^T(\mu,t) \geq \alpha\ \forall (u,t) \in A$. The proof of $\mu \geq 0$ is similar.

With (6) we conclude that $\mu t \leq \alpha\ \forall t \leq p^* \implies \mu p^* \leq \alpha$ together with (5) we find

$$\forall x \in D:\ \sum_{i=1}^{m} \lambda_i f_i(x) + \mu f_0(x) \geq \alpha \geq \mu p^*.$$

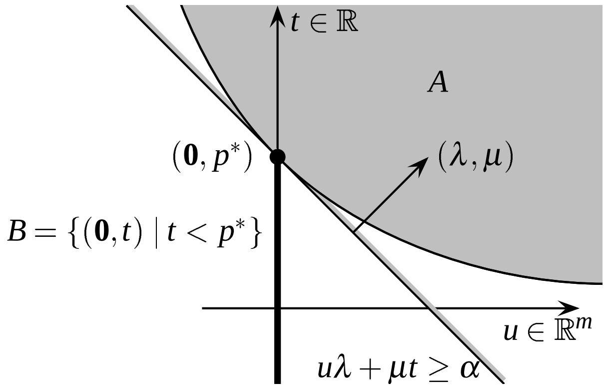

Figure 1: Situation in the proof of Theorem 10.

There are two different cases to analyse, the first one is when $\mu > 0$. Then

$$\sum_{i=1}^{m} \frac{\lambda_i}{\mu} f_i(x) + f_0(x) \geq p^* \implies g\left(\frac{\lambda}{\mu}\right) \geq p^*.$$

The second case is when $\mu = 0$. Then we have

$$\sum_{i=1}^{m} \lambda_i f_i(x) \geq 0 \quad \forall x \in D.$$

In particular for $\bar{x}$ (from Slater's condition) we obtain

$$\sum_{i=1}^{m} \lambda_i \underbrace{f_i(\bar{x})}_{<0} \geq 0,$$

this implies $\lambda = 0$, which is a contradiction.

□

References

[1] Boyd and Vandenberghe. *Convex Optimisation*. 2004.

[2] J. Matousek. *Lectures on discrete geometry*, volume 212 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2002.