1 Preliminaries

1.1 Σ is semi-positive definite

We can tell that the covariance matrix Σ is semi-positive definite since, for every vector $x \in \mathbb{R}^n$:

$$\begin{aligned} x^T \Sigma x &= x^T \left(\frac{1}{N} \sum_{i=1}^{N} (D_i - \overline{D})(D_i - \overline{D})^T \right) x \\ &= \frac{1}{N} \sum_{i=1}^{N} x^T (D_i - \overline{D})(D_i - \overline{D})^T x \\ &= \frac{1}{N} \sum_{i=1}^{N} \left[x^T (D_i - \overline{D}) \right]^2 \\ &\geq 0 \end{aligned} \tag{1}$$

For the sake of simplicity, we will consider for rest of the assignment that Σ is positive definite (i.e. the matrix is non-singular).

1.2 Convexity and solutions of the problem

Convexity

We know that

$$\begin{bmatrix} \max\limits_{x \in \mathbb{R}^n} \quad c(x) = R^T x & \\ \text{s.t.} \quad f(x) = \sum_{i=1}^{n} x_i - B = 0 & \\ \quad g_i(x) = x_i \geq 0 \qquad \text{for } i = 1, ..., n \end{bmatrix} = - \begin{bmatrix} \min\limits_{x \in \mathbb{R}^n} \quad \tilde{c}(x) = -c(x) = -R^T x & \\ \text{s.t.} \quad f(x) = \sum_{i=1}^{n} x_i - B = 0 & \\ \quad g_i(x) = x_i \geq 0 \qquad \text{for } i = 1, ..., n \end{bmatrix} \tag{2}$$

Since then, we know that the minimization problem in equation (2) is equivalent to the initial maximization problem. This problem is convex since

- It is a minimization problem ;
- The objective function $\tilde{c}$ is linear ;
- Equality constraint function f is linear ;
- Inequality constraint functions $\{g_i\}_{i=1}^n$ are linear.

Solutions

The solutions are easy to find : we look at the set $I = \{i \text{ s.t. } R_i \geq R_j \quad \forall j = 1, \ldots, n\}$. The set of solutions is thus $S = \{x \in \mathbb{R}^n \text{ s.t. } \sum_{i \in I} x_i = B \text{ and } x_j \geq 0 \quad \forall j = 1, \ldots, n\}$. In other terms, we look at the assets for which the expected return is the highest, and we invest all the budget B in these assets. Note that if $|I| = 1$, then $|S| = 1$ and, if $|I| > 1$, then $|S| = \infty$.

For any $x^* \in S$, we find the value of $c(x^*)$ on the diagram given in the statement in the following way. First, we compute the risk $r = x^{*T} \Sigma x^*$. Then the solution is the y-axis value of the intersection of the red line with the line where the x-axis is equal to r.

An example is given on Figure 1 for a x^* such that $r = x^{*T} \Sigma x^* = 0.6$.

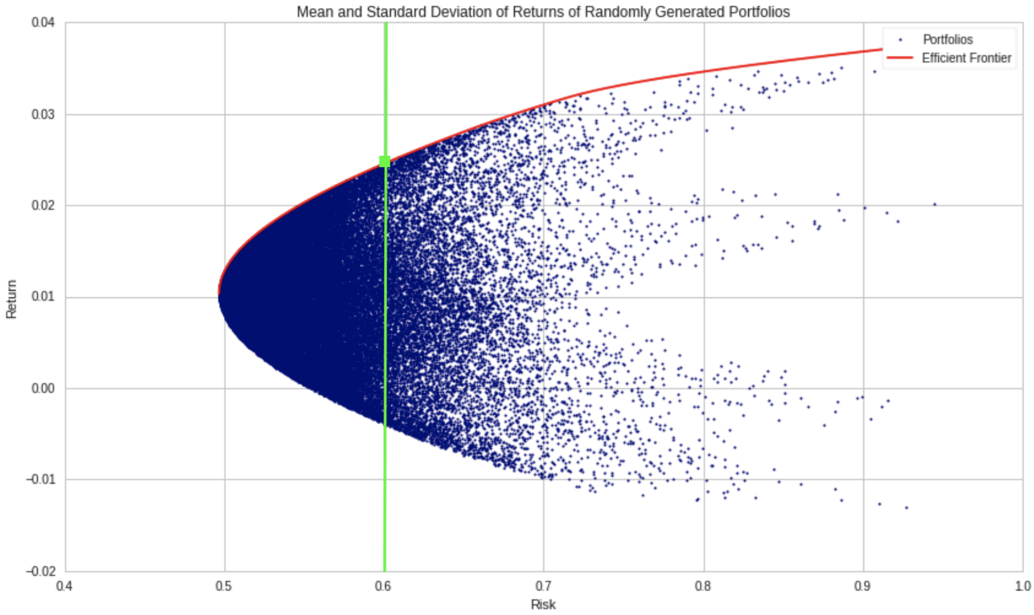

Figure 1: Application for $r = 0.6$

2 Risk-variance trade-off

2.1 Convexity and solutions of the problem

Convexity

We write the optimization problem as

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & c(x) = x^T \Sigma x \\ \text{s.t.} \quad & f_1(x) = \sum_{i=1}^{n} x_i - B = 0 \\ & f_2(x) = R^T x - T = 0 \\ & g_i(x) = x_i \geq 0 \qquad \text{for } i = 1, ..., n \end{aligned} \tag{3}$$

This problem is convex since

- It is a minimization problem ;
- The objective function c is quadratic with $\Sigma \succ 0$;
- Equality constraint function f_1 and f_2 are linear ;
- Inequality constraint functions $\{g_i\}_{i=1}^n$ are linear.

Solutions

For x^* an optimal solution of the problem (3), we find the value of $c(x^*)$ on the diagram given in the statement by looking at the x-axis value of the intersection of the red line with the line where the y-axis is equal to T.

An example is given on Figure 2 for $T = 0.03$.

Solution of such an optimization problem is unique since the objective function c is a quadratic form with Σ positive definite, i.e. a **strictly convex** function.

2.2 Parabolic border

The parabolic border represents :

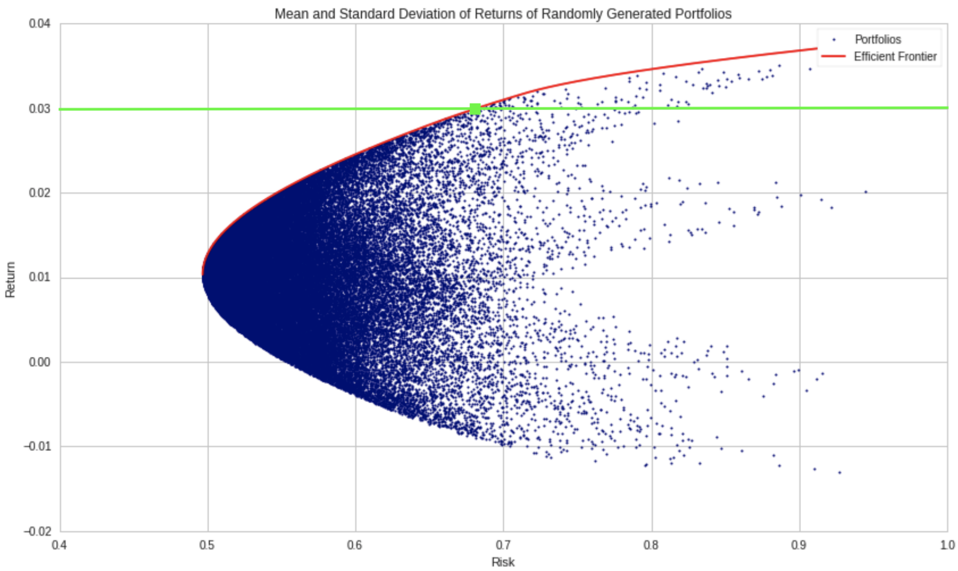

Figure 2: Application for $T = 0.03$

- For each possible risk : the maximal returns or,
- For each possible returns : the minimal risk.

In order to derive the parabolic border the we can see on Figures 1 and 2, we will relax the inequality constraints and solve the convex optimization problem (3) using the Lagrangian multipliers. We consider the Lagrangian function $\mathcal{L}(x, \lambda_1, \lambda_2) = c(x) + \lambda_1 f_1(x) + \lambda_2 f_2(x)$. The aim is to find $(x^*, \lambda_1^*, \lambda_2^*)$ such that

$$\nabla\mathcal{L}(x^*, \lambda_1^*, \lambda_2^*) = 0 \iff \nabla c(x^*) = -\lambda_1^* f_1(x^*) - \lambda_2^* f_2(x^*) \tag{4}$$

We redefine $f_1(x) = \sum_{i=1}^n x_i - B = u^T x - B$ with $u = (1\ 1\ \dots\ 1)^T$. Since $\nabla c(x) = 2\Sigma x$, $\nabla f_1(x) = u$ and $\nabla f_2(x) = R$, we have

$$2\Sigma x^* = -\lambda_1^* u - \lambda_2^* R \tag{5}$$

and, using the fact that Σ is non-singular

$$x^* = -\Sigma^{-1}\left(\frac{\lambda_1^*}{2}u + \frac{\lambda_2^*}{2}R\right) \tag{6}$$

Now we inject the optimal value of x^* in the constraints. We evaluate $f_1(x^*) = 0$ and $f_2(x^*) = 0$, which gives us the linear system

$$\begin{cases} (R^T\Sigma^{-1}R)\lambda_1^* + (R^T\Sigma^{-1}u)\lambda_2^* = -2T \\ (u^T\Sigma^{-1}R)\lambda_1^* + (u^T\Sigma^{-1}u)\lambda_2^* = -2B \end{cases} \tag{7}$$

We define $\alpha_1 = (R^T\Sigma^{-1}R)$, $\alpha_2 = (R^T\Sigma^{-1}u)$, $\alpha_3 = (u^T\Sigma^{-1}R) = \alpha_2$ (since Σ^{-1} is symmetric) and $\alpha_4 = (u^T\Sigma^{-1}u)$. Now equation (8) becomes

$$\begin{cases} \alpha_1\lambda_1^* + \alpha_2\lambda_2^* = -2T \\ \alpha_2\lambda_1^* + \alpha_4\lambda_2^* = -2B \end{cases} \tag{8}$$

The solution of this system is

$$\begin{aligned} \frac{\lambda_1^*}{2} &= -\frac{\alpha_4 T - \alpha_2 B}{\alpha_1\alpha_4 - \alpha_2^2} \\ \frac{\lambda_2^*}{2} &= -\frac{\alpha_1 B - \alpha_2 T}{\alpha_1\alpha_4 - \alpha_2^2} \end{aligned} \tag{9}$$

Now that λ_1^* and λ_2^* are computed, we can check at the value of the cost function $c(x^*)$ for the optimal solution (given in equation (6)). Remembering that $f_1(x^*) = 0$ and $f_2(x^*) = 0$, we get

$$\begin{aligned} c(x^*) &= x^{*T} \Sigma x^* \\ &= -x^{*T} \left(\frac{\lambda_1^*}{2} u + \frac{\lambda_2^*}{2} R \right) \\ &= -\frac{\lambda_1^*}{2} x^{*T} u - \frac{\lambda_2^*}{2} x^{*T} R \\ &= -\frac{\lambda_1^*}{2} B - \frac{\lambda_2^*}{2} T \\ &= \frac{\alpha_4 T^2 - 2\alpha_2 BT + \alpha_1 B^2}{\alpha_1 \alpha_4 - \alpha_2^2} \end{aligned} \tag{10}$$

Now, if we consider T as a variable, we can express the cost function as $\tilde{c}$, a function of T. $\tilde{c}$ is a parabola since

$$\tilde{c}(T) = \left(\frac{\alpha_4}{\alpha_1 \alpha_4 - \alpha_2^2} \right) T^2 - \left(\frac{2\alpha_2 B}{\alpha_1 \alpha_4 - \alpha_2^2} \right) T + \left(\frac{\alpha_1 B^2}{\alpha_1 \alpha_4 - \alpha_2^2} \right) \tag{11}$$

3 Markowitz model

We will now consider the maximization of a linear combination of the risk and overall return. Given a positive parameter μ, it consists in minimizing the objective $x^\top \Sigma x - \mu R^\top x$

3.1 Convexity and solutions of the problem

Convexity

The optimization problem is thus :

$$\begin{aligned} \min_x \quad & x^\top \Sigma x - \mu R^\top x \\ \text{s.t.} \quad & \sum_{i=1}^n x_i = B \\ & x_i \geq 0 \qquad \text{for } i = 1, ..., n \end{aligned} \tag{12}$$

We already proved the feasible region is convex in Question 1.2. Since it is a minimization problem, it remains to be proved that the objective function is convex :

- the quadratic function $x^\top \Sigma x$ is convex considering its hessian matrix is Σ and is positive definite.
- $R^\top x$ is an linear function, thus $-\mu R^\top x$ is also linear and, consequently, convex.
- the sum of two convex functions is also convex.

It leads to the conclusion that the problem is convex.

Solutions

Given the parameter μ, we can find the solutions of the problem on the diagram by finding the line of slope $1/\mu$ that is tangent to the efficient frontier of the diagram.

Indeed, the optimal value is given by $f(x^*) = x^{*\top} \Sigma x^* - \mu(R^\top x^*)$ which leads to a line of equation $Y = aX + b$ where $Y = R^\top x$, $a = \frac{1}{\mu}$, $X = x^\top \Sigma x$ and $b = -\frac{1}{\mu} f(x^*)$.

On Figure 3, for $\mu = 5$, we approximately find the line of equation $y = 0.2x - 0.006$ and $f(x^*) = 0.03$.

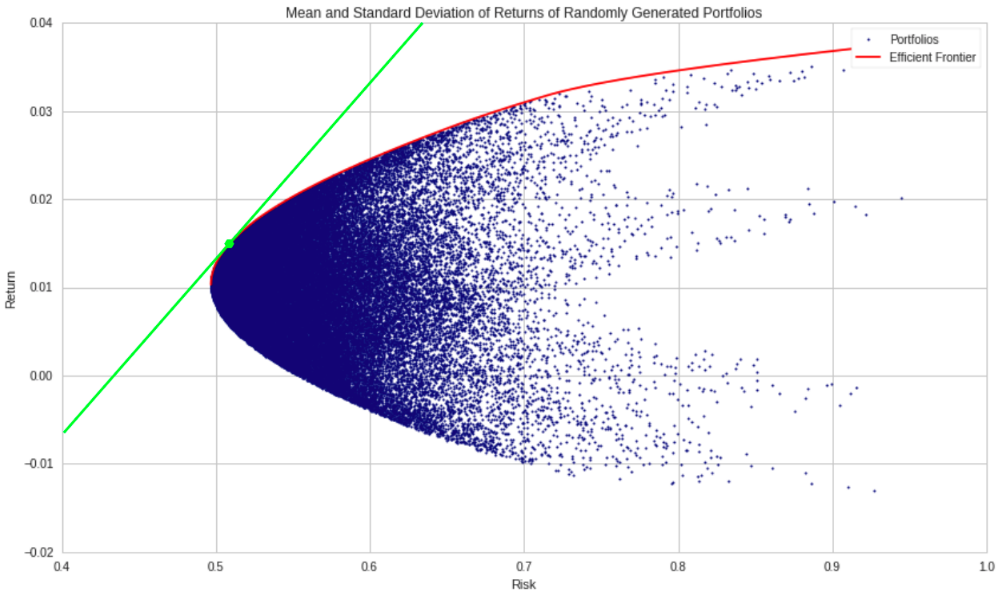

Figure 3: Application for $\mu = 5$

3.2 Lagrangian dual problem

Let be $f_0(x) = x^\top \Sigma x - \mu R^\top x$, the objective function of the primal function, $f_i(x) = -x_i \leq 0$ for $i = 1, ..., n$, the n inequality constraints and $h_1(x) = \sum_{i=1}^n x_i - B = 0$, the equality constraint.

The Lagrangian function is equal to :

$$\begin{aligned} \mathcal{L}(x, \lambda, \nu) &= f_o(x) + \sum_{i=1}^n \lambda_i f_i(x) + \nu_1 h_1(x) \\ &= x^\top \Sigma x - \mu R^\top x + \sum_{i=1}^n \lambda_i e_i^\top x + \nu_1 (u^\top x - B) \end{aligned} \tag{13}$$

where e_i is the i-th unit vector of the standard basis and u is the vector of all entries equals to one.

The dual function $g(\lambda, \nu)$ is equal to $\min_x \mathcal{L}(x, \lambda, \nu)$. We need to compute the partial derivative of the lagrangian function $\mathcal{L}(x, \lambda, \nu)$ with respect to the variable x :

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x}(x, \lambda, \nu) &= x^\top (\Sigma + \Sigma^\top) - \mu R^\top + \sum_{i=1}^n \lambda_i e_i^\top + \nu_1 u^\top \\ &= 2x^\top \Sigma - \mu R^\top + \sum_{i=1}^n \lambda_i e_i^\top + \nu_1 u^\top \end{aligned} \tag{14}$$

We have to find x^* such that the partial derivative of the lagrangian function $\mathcal{L}(x, \lambda, \nu)$ with respect to the variable x is equal to zero :

$$\frac{\partial \mathcal{L}}{\partial x}(x^*, \lambda, \nu) = 0 \iff x^* = \frac{1}{2} \Sigma^{-1} \left(\mu R - \sum_{i=1}^n \lambda_i e_i - \nu_1 u \right) \tag{15}$$

We obtain thus the following expression for the dual function :

$$
\begin{aligned}
g(\lambda,\nu) &= \min_x \mathcal{L}(x,\lambda,\nu) \\
&= \mathcal{L}(x^*,\lambda,\nu) \\
&= x^{*\top}\Sigma x^* - \mu R^\top x^* + \sum_{i=1}^n \lambda_i e_i^\top x^* + \nu_1(u^\top x^* - B) \\
&= \left(x^{*\top}\Sigma - \mu R^\top + \sum_{i=1}^n \lambda_i e_i^\top + \nu_1(u^\top) \right) x^* - \nu_1 B \\
&= \left(\frac{1}{2}\left(\mu R^\top - \sum_{i=1}^n \lambda_i e_i^\top - \nu_1 u^\top \right) - \mu R^\top + \sum_{i=1}^n \lambda_i e_i^\top + \nu_1 u^\top \right) \frac{1}{2}\Sigma^{-1} \left(\mu R - \sum_{i=1}^n \lambda_i e_i - \nu_1 u \right) - \nu_1 B \\
&= -\frac{1}{4}\left(\mu R^\top - \sum_{i=1}^n \lambda_i e_i^\top - \nu_1 u^\top \right) \Sigma^{-1} \left(\mu R - \sum_{i=1}^n \lambda_i e_i - \nu_1 u \right) - \nu_1 B \\
&= -\frac{\mu^2}{4} R^\top \Sigma^{-1} R + \frac{\mu}{2} R^\top \Sigma^{-1} (\sum_{i=1}^n \lambda_i e_i + \nu_1 u) - \frac{1}{4} (\sum_{i=1}^n \lambda_i e_i^\top + \nu_1 u^\top) \Sigma^{-1} (\sum_{i=1}^n \lambda_i e_i + \nu_1 u) - \nu_1 B
\end{aligned} \tag{16}
$$

We would like to have a concave function $g(\lambda_1, ..., \lambda_n, \nu_1)$ and this can be verified by computing the Hessian matrix of g based on the term $-\frac{1}{4}(\sum_{i=1}^n \lambda_i e_i^\top + \nu_1 u^\top)\Sigma^{-1}(\sum_{i=1}^n \lambda_i e_i + \nu_1 u)$:

$$
Hg = -\frac{1}{4}\begin{bmatrix} \Sigma^{-1} & \Sigma^{-1}u \\ u^\top \Sigma^{-1} & u^\top \Sigma^{-1} u \end{bmatrix} \tag{17}
$$

We already know that Σ is positive definite and thus Σ^{-1} also. We see that the last row of the matrix is a linear combination of the n other rows thus the eigenvalues of the Hessian matrix are the eigenvalues of Σ^{-1} (all strictly positive by assumption) multiplied by $\frac{1}{4}$ and the value zero.

Thus, the eigenvalues of the Hessian matrix of $g(\lambda_1, ..., \lambda_n, \nu_1)$ are all negative and by definition, the function g is concave. The dual function is useful because it provides a lower bound to our initial problem by solving the following optimization problem :

$$
\begin{aligned}
& \max_{\lambda,\nu_1} \quad g(\lambda,\nu_1) \\
& \text{s.t.} \quad \lambda_i \geq 0 \qquad \text{for } i = 1, ..., n \\
\iff & -\min_{\lambda,\nu_1} \quad -g(\lambda,\nu_1) \\
& \text{s.t.} \quad \lambda_i \geq 0 \qquad \text{for } i = 1, ..., n
\end{aligned} \tag{18}
$$

4 Sharpe ratio

We will now consider an other way of scalarizing the bi-criteria portfolio optimization problem. Given a independently available risk-free return $r > 0$, we will now maximize the difference between the overall return and this risk-free return, divided by the standard deviation of the portfolio's return.

This leads to an objective equal to $\frac{R^\top x - r}{\sqrt{x^\top \Sigma x}}$.

4.1 Convexity of the problem

The optimization problem is thus :

$$
\begin{aligned}
& -\min_x \quad \frac{r - R^\top x}{\sqrt{x^\top \Sigma x}} \\
& \text{s.t.} \quad \sum_{i=1}^n x_i = B \\
& \qquad x_i \geq 0 \qquad \text{for } i = 1, ..., n
\end{aligned} \tag{19}
$$

We already proved the feasible region is convex in Question 1.2. We want now to know if $F(x) = \frac{r - R^\top x}{\sqrt{x^\top \Sigma x}}$ is convex. We reformulate the function in the following way :

$$\begin{aligned} F(x) &= \frac{f_N(x)}{f_D(x)} \\ \text{with} \quad f_N(x) &= r - R^\top x \quad \text{linear function : } \lambda f_D(x) + (1-\lambda) f_D(y) = f_D(\lambda x + (1-\lambda) y) \\ f_D(x) &= \sqrt{x^\top \Sigma x} \quad \text{convex function : } \lambda f_D(x) + (1-\lambda) f_D(y) \geq f_D(\lambda x + (1-\lambda) y) \end{aligned} \tag{20}$$

By definition, the convexity condition is the following :

$$\begin{aligned} & \lambda F(x) + (1-\lambda) F(y) \geq F(\lambda x + (1-\lambda) y) \\ \iff & \lambda \frac{f_N(x)}{f_D(x)} + (1-\lambda) \frac{f_N(y)}{f_D(y)} \geq \frac{f_N(\lambda x + (1-\lambda) y)}{f_D(\lambda x + (1-\lambda) y)} \\ \iff & \frac{\lambda f_N(x) f_D(y) + (1-\lambda) f_N(y) f_D(x)}{f_D(x) f_D(y)} \geq \frac{f_N(\lambda x + (1-\lambda) y)}{f_D(\lambda x + (1-\lambda) y)} \\ \iff & [\lambda f_N(x) f_D(y) + (1-\lambda) f_N(y) f_D(x)] f_D(\lambda x + (1-\lambda) y) \geq f_N(\lambda x + (1-\lambda) y) f_D(x) f_D(y) \end{aligned} \tag{21}$$

Using properties that f_D is convex and f_N is linear, we obtain :

$$[\lambda f_N(x) f_D(y) + (1-\lambda) f_N(y) f_D(x)][\lambda f_D(x) + (1-\lambda) f_D(y)] \geq [\lambda f_N(x) + (1-\lambda) f_N(y)] f_D(x) f_D(y) \tag{22}$$

By factorization and division by $\lambda(1-\lambda)$, it leads to the convexity condition :

$$\begin{aligned} & -f_N(x) f_D(x) f_D(y) + f_N(y) f_D(x)^2 + f_N(x) f_D(y)^2 - f_N(y) f_D(y) f_D(x) \geq 0 \\ \iff & [f_D(x) - f_D(y)][f_N(y) f_D(x) - f_N(x) f_D(y)] \geq 0 \end{aligned} \tag{23}$$

In other words, we have the following conditions :

$$\begin{aligned} & f_D(x) \leq f_D(y) \quad \text{and} \quad f_N(y) f_D(x) \leq f_N(x) f_D(y) \\ \text{or} & \\ & f_D(y) \leq f_D(x) \quad \text{and} \quad f_N(x) f_D(y) \leq f_N(y) f_D(x) \end{aligned} \tag{24}$$

For example, we can take $x = [B, 0, 0, ..., 0]^\top$ and $y = [0, B, 0, ..., 0]^\top$, convexity conditions become :

$$\begin{aligned} & \sqrt{B \Sigma_{11} B} \leq \sqrt{B \Sigma_{22} B} \quad \text{and} \quad (r - R_2 B) \sqrt{B \Sigma_{11} B} \leq (r - R_1 B) \sqrt{B \Sigma_{22} B} \\ \text{or} & \\ & \sqrt{B \Sigma_{22} B} \leq \sqrt{B \Sigma_{11} B} \quad \text{and} \quad (r - R_1 B) \sqrt{B \Sigma_{22} B} \leq (r - R_2 B) \sqrt{B \Sigma_{11} B} \\ & \Sigma_{11} \leq \Sigma_{22} \quad \text{and} \quad R_2 \Sigma_{11} \geq R_1 \Sigma_{22} \\ \implies \text{or} & \\ & \Sigma_{22} \leq \Sigma_{11} \quad \text{and} \quad R_1 \Sigma_{22} \geq \Sigma_{11} R_2 \end{aligned} \tag{25}$$

Thus, there are conditions on vector R and matrix Σ and we can find some R and Σ such that above conditions are not respected and consequently, the function is not convex.

4.2 Convex reformulation of the problem

First, we can reformulate the numerator by using the equivalence $\sum_{i=1}^n x_i = B \iff \frac{1}{B} \sum_{i=1}^n x_i = 1$:

$$-F(x) = \frac{R^\top x - r}{\sqrt{x^\top \Sigma x}} = \frac{R^\top x - \frac{1}{B} u^\top x}{\sqrt{x^\top \Sigma x}} = \frac{\tilde{R}^\top x}{\sqrt{x^\top \Sigma x}} \quad \text{with } \tilde{R}_i = R_i - \frac{1}{B} \quad \text{for } i = 1, ..., n \tag{26}$$

We can prove that $F(x)$ is 0-homogeneous, i.e. $F(\lambda x) = F(x)$ for $\lambda > 0$ since :

$$F(\lambda x) = \frac{-\lambda \tilde{R}^\top x}{\sqrt{\lambda^2 x^\top \Sigma x}} = \frac{\lambda}{|\lambda|} \frac{-\tilde{R}^\top x}{\sqrt{x^\top \Sigma x}} = \frac{-\tilde{R}^\top x}{\sqrt{x^\top \Sigma x}} \quad \text{for } \lambda > 0 \tag{27}$$

We can now reformulate our problem in the following way :

$$\begin{aligned} \max_y \quad & \frac{1}{\sqrt{y^\top \Sigma y}} \\ \text{s.t.} \quad & \tilde{R}^\top y = 1 \\ & y_i \geq 0 \qquad \text{for } i = 1, ..., n \end{aligned} \tag{28}$$

Indeed, given an optimal solution x^* of the problem (19), since $F(x)$ is 0-homogeneous, x^* and λx^* ($\forall \lambda > 0$) are solutions of the problem without the condition $\sum_{j=1}^n x_j^* = B$. Thus, given an optimal solution y^* of the problem (28), if we find α such that $x^* = \alpha y^*$ and :

$$\sum_{j=1}^n x_j^* = B \quad \text{and} \quad x_j^* \geq 0 \quad \text{for } i = 1, ..., n \tag{29}$$

Then, y^* is also solution of the first problem because we have :

$$-F(x^*) = -F(\alpha y^*) = -F(y^*) = \frac{\tilde{R}^\top y^*}{\sqrt{y^{*\top} \Sigma y^*}} = \frac{1}{\sqrt{y^{*\top} \Sigma y^*}} \tag{30}$$

This can be verified with $x^* = \left(\frac{B}{\sum_{i=1}^n y_i^*} \right) y^*$:

$$\sum_{j=1}^n x_j^* = \sum_{j=1}^n \frac{B}{\sum_{i=1}^n y_i^*} y_j^* = B \frac{\sum_{j=1}^n y_j^*}{\sum_{i=1}^n y_i^*} = B \tag{31}$$

And

$$\begin{aligned} & y_i \geq 0 \qquad \text{for } i = 1, ..., n \qquad \text{and} \qquad B > 0 \\ \implies & y_i \geq 0 \qquad \text{for } i = 1, ..., n \qquad \text{and} \qquad \frac{B}{\sum_{i=1}^n y_i^*} > 0 \\ \implies & x_i \geq 0 \qquad \text{for } i = 1, ..., n \end{aligned} \tag{32}$$

Finally, the optimization problem from equation (28) is equivalent to :

$$\begin{aligned} \min_y \quad & \sqrt{y^\top \Sigma y} \\ \text{s.t.} \quad & \tilde{R}^\top y = 1 \\ & y_i \geq 0 \qquad \text{for } i = 1, ..., n \end{aligned} \tag{33}$$

Find the optimal solution x^* our initial problem from equation (19) is thus equivalent to find the optimal value y^* of problem (33) and compute $x^* = \left(\frac{B}{\sum_{i=1}^n y_i^*} \right) y^*$ with the advantage that our new formulation is a convex optimization problem.

Bonus : solutions

We can find the optimal value of our optimization $-\min_x F(x) = \frac{r - R^\top x}{\sqrt{x^\top \Sigma x}}$ by a similar way as in Question 3.1.

The optimal value is given by $-F(x^*)$ and we have $F(x^*) = \frac{r - R^\top x*}{\sqrt{x^{*\top} \Sigma x^*}} \iff R^\top x* = -F(x^*)\sqrt{x^{*\top} \Sigma x^*} + r$ which leads to a curve of equation $Y = a\sqrt{X} + b$ where $Y = R^\top x$, $a = -F(x^*)$, $X = x^\top \Sigma x$ and $b = r$.

In order to find the optimal objective, we search the parameter $a = -F(x^*)$ such that the curve $Y = a\sqrt{X} + r$ intersects the parabola (section 2.2) at one point.