The portfolio optimization problem

- An investor has a fixed amount of money to invest in a portfolio of \( n \) risky assets \( S^1, \ldots, S^n \), and a riskless asset \( S^0 \);
- We consider the portfolio’s return over a fixed investment period;
- The \textit{random} return of asset \( S^i \) over this period is \( R_i \).
- \( \mu = \mathbb{E}[R] \): vector of expected returns. Assumed known.
- \( Q = \text{Cov}(R) \): variance-covariance matrix of returns. Assumed known.

**Question**

*What proportion of wealth should the investor invest in asset \( S^i \) ?*
The portfolio optimization problem

• decision variables: \( x_i \): proportion of wealth invested in asset \( S^i \).

• constraints:
  • the entire wealth is assumed invested: \( \sum_{i=1}^{n} x_i = 1 \)
  • short-selling of assets is not allowed: \( x_i \geq 0 \)
  • bounds on exposure to groups of assets \( \sum_{i \in g} x_i \leq b \)

• objective function: The investor wants to maximize the expected return while minimizing risk.
  • Expected return: \( \mathbb{E}[R^T x] = \mu^T x \)
  • How do we measure risk?
The Konno & Yamazaki Model

Konno & Yamazaki: measure risk using the $\ell_1$ norm:

$$\ell(x) = \mathbb{E}[|(R - \mu)^T x|]$$

- This is a measure of volatility like the variance would be.
- Why preferring it over the variance $\sigma^2(x)$ of the portfolio return?

$$\sigma^2(x) = \mathbb{E}[((R - \mu)^T x)^2]$$

Two reasons:

1. If $R \sim \mathcal{N}(\mu, Q)$, then $\sigma^2(x) = x^T Q x$ and

$$\ell(x) = \sqrt{\frac{2\sigma^2(x)}{\pi}}$$

and minimizing $\ell(x)$ is equivalent to minimizing $\sigma^2(x)$. (The assumption is unrealistic)

2. The resulting optimization problem is linear rather than quadratic, so in practice we can handle many more assets.
A linear optimization model

- **objective function**: The investor wants to maximize the expected return while minimizing risk.

- Two objectives? Possible, but we are not going there.
- We fix a minimum level for expected return (or a maximum level of risk) and optimize for minimal risk (or maximum return)
  E.g., add constraint $r_{\text{min}} \leq \mu^T x$.
- More constraints can be added to limit short selling, sector allocation, transaction costs, and so on.
  Constraints: $Ax \geq a$, $Bx = b$ (including the above)
A linear optimization model

- **objective function:** The investor wants to maximize the expected return while minimizing risk.

- Two objectives? Possible, but we are not going there.
- We fix a minimum level for expected return (or a maximum level of risk) and optimize for minimal risk (or maximum return).
  
  E.g., add constraint $r_{\text{min}} \leq \mu^T x$.

- More constraints can be added to limit short selling, sector allocation, transaction costs, and so on.

  Constraints: $Ax \geq a$, $Bx = b$ (including the above)

Konno & Yamazaki model

\[
(KY) \quad \min_x \ell(x) \\
\text{s.t. } Ax \geq a \\
Bx = b
\]
Estimating parameters

- The expectation \( \ell(x) = E[|| (R - \mu)^T x ||] \) may not be computable with only partial information (i.e., without knowing the distribution of \( R \)).
- We can replace it with an empirical estimate.
- For each investment \( S^i \) let \( \{r_{it}, t = 1, \ldots, T\} \) be a collection of historic returns over period of times of the same length.
- We can replace \( \mu_i \) and \( \ell(x) \) with their empirical estimates:

\[
\hat{\mu}_i = \frac{1}{T} \sum_{t=1}^{T} r_{it}
\]

\[
\hat{\ell}(x) = \frac{1}{T} \sum_{t=1}^{T} \left| \sum_{i=0}^{n} (r_{it} - \hat{\mu}_i) x_i \right|
\]
The revised model

Using these replacements, the problem \((KY)\) becomes

\[
(KY') \quad \min_x \frac{1}{T} \sum_{t=1}^{T} \left| \sum_{i=0}^{n} (r_{it} - \hat{\mu}_i)x_i \right|
\]

s.t. \(Ax \geq a\)

\(Bx = b\)

The optimization function is no longer linear (absolute value)

**Question**

*How can we transform it into a LP?*

*(Hint: replace \(|w|\) with a function of two new variables \(y, z \geq 0\), and add a constraint involving \(w, y, z\))*
The revised model

Using these replacements, the problem \((KY)\) becomes

\[
(KY') \quad \min_x \frac{1}{T} \sum_{t=1}^{T} \left| \sum_{i=0}^{n} (r_{it} - \hat{\mu}_i) x_i \right|
\]

s.t. \(Ax \geq a\)

\(Bx = b\)

The optimization function is no longer linear (absolute value)

**Question**

*How can we transform it into a LP?*

*(Hint: replace \(|w|\) with a function of two new variables \(y, z \geq 0\), and add a constraint involving \(w, y, z\))*

\[
(KY'') \quad \min_{x, y, z} \frac{1}{T} \sum_{t=1}^{T} (y_t + z_t)
\]

s.t. \(y_t - z_t = \sum_{i=0}^{n} (r_{it} - \hat{\mu}_i) x_i\)
Markowitz Portfolio Optimization

- In the traditional Markowitz formulation we measure the risk using the variance of the portfolio return:

\[ \sigma^2(R^T x) = x^T Q x \]

- The optimization problem becomes:

\[
\begin{align*}
\min & \quad x^T Q x \\
\text{s.t.} & \quad m^T x \geq r_{\min} \\
& \quad 1^T x = 1
\end{align*}
\]
Markowitz Portfolio Optimization

Variants:

Maximize the risk-adjusted return:

$$\max \mu^T x - \delta x^T Q x$$

subject to

$$Ax \geq a$$

$$Bx = b$$

where $\delta$ is a risk-aversion constant.

Maximize the return given an acceptable level of risk:

$$\max \mu^T x$$

subject to

$$x^T Q x \leq \sigma_{max}^2$$

$$Ax \geq a$$

$$Bx = b$$
In practice the optimal portfolio weights $x^*$ computed by the above models have undesirable features:

- the weights $x_i$ may be very large for a few of the assets, which means taking large short/long positions in a small number of assets
- the weights are instable: a small error in the estimation of $\mu$ and $Q$ leads to big changes in $x^*$
- There is no bound in transaction costs, when considering multiple periods of time
- We end up with a portfolio whose true risk we don’t really know. All these problem can be mitigated/solved by adding additional constraints
**Additional constraints**

- To avoid excessive short or long position we can use box constraints:
  \[ l_i \leq x_i \leq u_i \]

- To avoid large transaction costs, we can impose an upper bound \( h \) on the proportion of the portfolio value that is traded. Let \( \tilde{x}_i \) be the old proportion of asset \( S^i \) and \( x_i \) be the new proportion.
  \[ x_i - \tilde{x}_i \leq y_i, \quad y_i \geq 0 \]
  \[ \tilde{x}_i - x_i \leq z_i, \quad z_i \geq 0 \]
  \[ \sum_{i=0}^{n} (y_i + z_i) \leq h \]

- To ensure diversification, we can group assets into possibly overlapping classes \( G_1, \ldots, G_k \), and impose constraints on the proportion of wealth allocated to each class:
  \[ l_j \leq \sum_{i \in G_j} x_i \leq u_j \]
The efficient frontier

- Assumption: $Q$ is positive definite, no risk-free asset.
- Let

$$\sigma_{\text{min}}^2 = \min_x x^T Q x$$

s.t. $A x \geq a$

$$B x = b$$

- Since $Q$ is positive definite, we have $\sigma_{\text{min}} > 0$
- Let

$$(R) \quad r(\sigma) = \max_x \mu^T x$$

s.t. $A x \geq a$

$$B x = b$$

$$x^T Q x \leq \sigma^2$$

The function $r(\sigma)$ is well define for $\sigma \geq \sigma_{\text{min}}$, as $(R)$ has feasible solutions.
The efficient frontier

- We have $\mu^T x \leq r(\sqrt{x^T Q x})$ for all feasible $x$
- It never makes sense to hold a portfolio $x$ for which

$$\mu^T x < r(\sqrt{x^T Q x})$$

since the portfolio $x^*$ obtained by solving $R$ with $\sigma^2 = x^T Q x$ would yield the higher return

$$\mu^T x^* = r(\sqrt{x^T Q x})$$
The efficient frontier

- We have $\mu^T x \leq r(\sqrt{x^T Q x})$ for all feasible $x$
- It never makes sense to hold a portfolio $x$ for which
  \[ \mu^T x < r(\sqrt{x^T Q x}) \]
  since the portfolio $x^*$ obtained by solving $R$ with $\sigma^2 = x^T Q x$
  would yield the higher return
  \[ \mu^T x^* = r(\sqrt{x^T Q x}) \]

Definition (Efficient portfolios)

Portfolios that satisfy the relation
\[ \mu^T x^* = r(\sqrt{x^T Q x}) \]
are called efficient. The curve $\sigma \rightarrow r(\sigma)$, defined for $\sigma \geq \sigma_{\text{min}}$ is called the efficient frontier.

Fact

*The efficient frontier is a concave function.*
• An alternative view of the same relation as \( r(\sigma) \) is to consider, any \( r \geq r_{\text{min}} = r(\sigma_{\text{min}}) \) and compute

\[
\sigma^2(r) = \min_x x^T Q x
\]

s.t. \( Ax \geq a \)

\( Bx = b \)

\( \mu^T x \geq r \)

• Using optimality conditions one can show that the relations \( \sigma(r) \) and \( r(\sigma) \) are inverses of each other.
We now consider the situation where the assets are
• One risk-free asset $S^0$ with return $r_f$; and
• $n$ risky assets $S^1, \ldots, S^n$, with random return vector $R$ with
  $\mathbb{E}[R] = \mu$ and $\text{Cov}(R) = Q$ positive definite.
• Let

$\tilde{x} = \begin{bmatrix} x_0 \\ x \end{bmatrix}, \tilde{R} = \begin{bmatrix} r_f \\ R \end{bmatrix}, \tilde{\mu} = \begin{bmatrix} r_f \\ \mu \end{bmatrix}, \tilde{Q} = \begin{bmatrix} 0 & 0 \\ 0 & Q \end{bmatrix}$

where $\tilde{Q}$ is the covariance matrix of $\tilde{R}$. 
Consider the Markowitz problem

\[(P) \quad \max_{\tilde{x}} \tilde{x}^T \tilde{Q} \tilde{x} \]

\[\text{s.t. } \tilde{A} \tilde{x} \geq \tilde{a} \]
\[\tilde{B} \tilde{x} \geq b \]
\[\tilde{\mu}^T \tilde{x} \geq \tilde{r} \]

where the constraint matrices are of the form

\[\tilde{A} = \begin{bmatrix} 0 & e \\ a & A \end{bmatrix}, \tilde{B} = [bB], \quad \tilde{a} = \begin{bmatrix} 0 \\ a \end{bmatrix} \]

where \( e = [1 \cdots 1] \) and \( \tilde{B} \) has first row \([1e]\).
• Let $e_0 = [10]^T$ be the portfolio that corresponds to investing the entire wealth in the risk-free asset. $e_0$ is feasible.

• The constraint $\tilde{B}\tilde{x}$ contains the constraint $\sum_i x_i = 1$.

• The constrain structure implies that, for all $x \in \mathbb{R}^n$ s.t. $Ax \geq a$ and $Bx = b$, and for all $\theta \leq 1$, the portfolio

$$\begin{pmatrix}
\theta e_0 \\
(1 - \theta)x
\end{pmatrix}$$

is feasible for $(P)$
Markowitz Model without Risk-free asset

• Now consider the associated Markowitz problem

\[
(M) \quad \max_x x^T Q x \\
\text{s.t. } Ax \geq a \\
Bx \geq b \\
\mu^T x \geq r
\]

obtained by only considering risky assets

• Let \( r(\sigma), \sigma(r), r_{\min}, \) and \( \sigma_{\min} \) be as introduced earlier, and let \( \tilde{r}(\tilde{\sigma}), \tilde{\sigma}(\tilde{r}), \tilde{r}_{\min}, \) and \( \tilde{\sigma}_{\min} \) be the corresponding objects for \((P)\).

• We have \( \tilde{\sigma}_{\min} = 0 \) and \( \tilde{r}_{\min} = r_f \).

• We assume that \((M)\) is feasible for some \( r > r_f \), otherwise investing a positive amount in risky assets is pointless.
The efficient frontier

Theorem

There exists a portfolio $x^m \in \mathbb{R}^n$ on the efficient frontier of $(M)$ such that the efficient frontier of $(P)$ is given by the ray $\{\tilde{x}(\theta) : \theta \leq 1\}$, where

$$
\tilde{x}(\theta) = \begin{pmatrix}
\theta e_0 \\
(1 - \theta)x^m
\end{pmatrix}
$$

- In other words, for any $\tilde{\sigma}^2 \geq 0$, there exits a $\theta \leq 1$ such that the portfolio $\tilde{x}(\theta)$ achieves the maximum return $\tilde{\mu}^T\tilde{x}(\theta) = \tilde{r}(\tilde{\sigma})$ at the risk level $\tilde{\sigma}^2$.
- This means that the relative proportions of wealth allocation among the risky assets is the same for all investors, no matter how risk averse they are.
The maximum Sharpe ratio problem

- In the proof of the theorem, the portfolio $x^m$ is chosen as an optimal solution of the *maximum Sharpe ratio problem*

$$\text{(SR) } \max_x \frac{\mu^T x - r_f}{\sqrt{x^t Q x}}$$

s.t. $Ax \geq a$

$B_x = b$
The maximum Sharpe ratio problem

• In the proof of the theorem, the portfolio $x^m$ is chosen as an optimal solution of the maximum Sharpe ratio problem

\[
(SR) \quad \max_x \frac{\mu^T x - r_f}{\sqrt{x^t Q x}}
\]

s.t. $Ax \geq a$

$Bx = b$

• Since the first constraint in $Bx = b$ is $e^T x = 1$, we can rewrite $\mu^T x = r_f$ as $(\mu - r_f e)^T x$.

• Then, if $y = \tau x$ for some $\tau > 0$, then

\[
\frac{(\mu^T - r f e)x}{\sqrt{x^t Q x}} = \frac{(\mu^T - r f e)y}{\sqrt{y^t Q y}}
\]

• In particular, if $\mu^T x > r_f$, we may choose $\tau > 0$ such that

\[
(\mu - r_f e)^T y = 1
\]
Solving the maximum Sharpe ratio problem

We can then solve \((SR)\) by solving the convex QP:

\[
(HSR) \quad \min_{y, \tau} y^T Qy \\
\text{s.t. } (\mu - rf e)^T y = 1 \\
Ay \geq \tau a \\
By = \tau b \\
\tau \geq 0
\]

And we can obtain \(x^*\) as \(y^*/\tau^*\) (\(\tau\) is positive for any feasible solution)
The market portfolio

• The portfolio corresponding to $x^m$ is called the market portfolio, the ray $\{\tilde{x}(\theta) : \theta \leq 1\}$ the capital market line and the gradient

$$\frac{\mu^T x^m - r_f}{\sqrt{x^m^T Q x^m}}$$

the Maximum Sharpe Ratio

• The MSR has an interpretation of market price of risk

• An investor is willing to take on an additional unit of risk (measured in terms of the standard deviation of the portfolio return) only if the expected portfolio return increases by the maximum Sharpe ratio
A Quick Tour of Nonlinear Programming

• Recall the fundamental theorem of asset pricing from Lecture 6.

• Stripping away the context, this theorem really just says something like:

**Theorem 1.**

Let \( A \in \mathbb{R}^{m \times n} \) and \( c \in \mathbb{R}^m \). Then exactly one of the following systems has a solution:

I. \( Ad \geq 0 \) and \( c^\top d < 0 \).

II. \( A^\top y = c \) and \( y \geq 0 \).

• In this form, the theorem is known as the Farkas Theorem and can be seen as the basis for much of optimization theory.

• It is easy to see the relationship to LP duality and this is another way of deriving strong duality for LP.
Setting

- You are a portfolio manager;
- You would like to understand how your “friend” Sally, a portfolio manager build her portfolio, i.e., which mix of stocks goes into her portfolio, also known as the style of the portfolio;
- Sally is secretive on the mix, but likes to publish the returns of her portfolio over time;
- You also have access to the returns of index funds tracking different sectors of the market;
- (You would also like to understand how good she is as a portfolio manager;)

Definition (Return-Based Style Analysis (RBSA))

A technique using constrained optimization to determine the style of a portfolio using the return time series of the portfolio and of a number of other asset classes (factors).
### Returns-based style analysis

#### Setting
- You are a portfolio manager;
- You would like to understand how your “friend” Sally, a portfolio manager build her portfolio, i.e., which mix of stocks goes into her portfolio, also known as the *style* of the portfolio;
- Sally is secretive on the mix, but likes to publish the returns of her portfolio over time;
- You also have access to the returns of index funds tracking different sectors of the market;
- *(You would also like to understand how good she is as a portfolio manager;)*

#### Definition (Return-Based Style Analysis (RBSA))

A technique using *constrained optimization* to determine the style of a portfolio using the return time series of the portfolio and of a number of other asset classes (*factors*).
RBSA Mathematical Model

Fundamentally a linear model for regression.

### Data

- $R_t, t = 1, \ldots, T$: the return of Sally’s portfolio over $T$ fixed time intervals (e.g., $R_t$ is the monthly return, and $T = 12$ months);
- $F_{it}, i = 1, \ldots, n, t = 1, \ldots, T$: the returns of factor $i$ over $T$ fixed time intervals (same intervals as $R_t$);
RBSA Mathematical Model

Fundamentally a linear model for regression.

### Data

- \( R_t, \ t = 1, \ldots, T \): the return of Sally’s portfolio over \( T \) fixed time intervals (e.g., \( R_t \) is the monthly return, and \( T = 12 \) months);
- \( F_{it}, \ i = 1, \ldots, n, \ t = 1, \ldots, T \): the returns of factor \( i \) over \( T \) fixed time intervals (same intervals as \( R_t \));

### Model

\[
R_t = w_{1t}F_{1t} + w_{2t}F_{2t} + \cdots + w_{nt}F_{nt} + \varepsilon_t = F_t^T w_t + \varepsilon_t
\]

- \( w_{it} \): sensitivity of \( R_t \) to factor \( i \);
- \( \varepsilon_t \): non-factor return.
Interpretation

\[ R_t = w_{1t}F_{1t} + w_{2t}F_{2t} + \cdots + w_{nt}F_{nt} + \varepsilon_t = F_t^Tw_t + \varepsilon_t \]

- Assume the \( F_{it} \) are returns of passive investments (e.g., index funds);
\[ R_t = w_1 t F_1 t + w_2 t F_2 t + \cdots + w_n t F_n t + \varepsilon_t = F_t^T w_t + \varepsilon_t \]

- Assume the \( F_{it} \) are returns of passive investments (e.g., index funds);
- Then the term \( F_t^T w_t \) is the return of a benchmark portfolio of passive investments;
Interpretation

\[ R_t = w_1 t F_{1t} + w_2 t F_{2t} + \cdots + w_n t F_{nt} + \varepsilon_t = F_t^T w_t + \varepsilon_t \]

• Assume the \( F_{it} \) are returns of passive investments (e.g., index funds);
• Then the term \( F_t^T w_t \) is the return of a benchmark portfolio of passive investments;
• Then the term \( \varepsilon_t \) is the difference between the passive benchmark and the active strategy followed by Sally.
Interpretation

\[ R_t = w_1 F_{1t} + w_2 F_{2t} + \cdots + w_n F_{nt} + \varepsilon_t = F_t^T w_t + \varepsilon_t \]

- Assume the \( F_{it} \) are returns of passive investments (e.g., index funds);
- Then the term \( F_t^T w_t \) is the return of a benchmark portfolio of passive investments;
- Then the term \( \varepsilon_t \) is the difference between the passive benchmark and the active strategy followed by Sally.
- If the passive investments considered together are representative of the market, then \( \varepsilon_t \) measures the additional (or negative) return due to Sally’s ability as a portfolio manager.
Optimization problem

\[ R_t = w_{1t} F_{1t} + w_{2t} F_{2t} + \cdots + w_{nt} F_{nt} + \varepsilon_t = F_t^T w_t + \varepsilon_t \]

Additional assumptions/constraints:

- We assume that \( w_{it} = w_i \), i.e., the weights do not change over time.
Optimization problem

\[ R_t = w_{1t}F_{1t} + w_{2t}F_{2t} + \cdots + w_{nt}F_{nt} + \varepsilon_t = F_t^T w_t + \varepsilon_t \]

**Additional assumptions/constraints:**

- We assume that \( w_{it} = w_i \), i.e., the weights do not change over time.
- The weights should be non-negative, and should sum to one;
Optimization problem

\[ R_t = w_1 t F_{1t} + w_2 t F_{2t} + \cdots + w_n t F_{nt} + \varepsilon_t = F_t^T w_t + \varepsilon_t \]

Additional assumptions/constraints:

- We assume that \( w_{it} = w_i \), i.e., the weights do not change over time.
- The weights should be non-negative, and should sum to one;

These allow us to formulate the constraints of our optimization problem:

\[ \min \text{??} \]

\[ \text{s.t. } \sum_{i=1}^{n} w_i = 1 \]

\[ w_i \geq 0, i = 1, \ldots, n \]

What about the objective function? Any idea?
Optimization problem (cont.)

- If $\varepsilon_t$ measures Sally’s ability as a portfolio manager, we can assume that it is approximately constant over time.

\[
\min_{\mathbf{w}} \in \mathbb{R}^n \quad \text{Var}(\varepsilon_t) = \text{Var}(R_t - F_T^T \mathbf{w}) \\
\text{s.t.} \\
\forall i = 1, \ldots, n, \quad w_i = 1, \quad w_i \geq 0,
\]

Is the objective function convex in $\mathbf{w}$?
Optimization problem (cont.)

- If $\varepsilon_t$ measures Sally’s ability as a portfolio manager, we can assume that it is approximately constant over time.
- I.e., we want the plots of the returns of Sally’s portfolios and of the benchmark portfolios to be curves with approximately constant distance.
Optimization problem (cont.)

• If $\varepsilon_t$ measures Sally’s ability as a portfolio manager, we can assume that it is approximately constant over time.

• I.e., we want the plots of the returns of Sally’s portfolios and of the benchmark portfolios to be curves with *approximately constant distance*.

• I.e., we want $\varepsilon_t$ to have the smallest possible *variance* over time.
• If \( \varepsilon_t \) measures Sally’s ability as a portfolio manager, we can assume that it is approximately constant over time.

• I.e., we want the plots of the returns of Sally’s portfolios and of the benchmark portfolios to be curves with *approximately constant distance*.

• I.e., we want \( \varepsilon_t \) to have the smallest possible *variance* over time.

**Formulation**

\[
\begin{align*}
\min_{w \in \mathbb{R}^n} & \quad \text{Var}(e_t) = \text{Var}(R_t - F_t^T w) \\
\text{s.t.} & \quad \sum_{i=1}^{n} w_i = 1 \\
& \quad w_i \geq 0, \; i = 1, \ldots, n
\end{align*}
\]
Optimization problem (cont.)

• If $\varepsilon_t$ measures Sally’s ability as a portfolio manager, we can assume that it is approximately constant over time.

• I.e., we want the plots of the returns of Sally’s portfolios and of the benchmark portfolios to be curves with *approximately constant distance*.

• I.e., we want $\varepsilon_t$ to have the smallest possible *variance* over time.

**Formulation**

\[
\min_{w \in \mathbb{R}^n} \text{Var}(e_t) = \text{Var}(R_t - F_t^T w)
\]

s.t. \[\sum_{i=1}^{n} w_i = 1\]

\[w_i \geq 0, \; i = 1, \ldots, n\]

Is the objective function convex in $w$?
Objective function

Let

\[ R = \begin{bmatrix} R_1 \\ \vdots \\ R_T \end{bmatrix}, \quad \text{and} \quad F = \begin{bmatrix} F_1^T \\ \vdots \\ F_T^T \end{bmatrix}, \quad \text{and} \quad e = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \]
Objective function

Let

\[ R = \begin{bmatrix} R_1 \\ \vdots \\ R_T \end{bmatrix}, \text{ and } F = \begin{bmatrix} F^T_1 \\ \vdots \\ F^T_T \end{bmatrix}, \text{ and } e = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \]

We have
Objective function

Let

\[ R = \begin{bmatrix} R_1 \\ \vdots \\ R_T \end{bmatrix}, \quad \text{and} \quad F = \begin{bmatrix} F_1^T \\ \vdots \\ F_T^T \end{bmatrix}, \quad \text{and} \quad e = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \]

We have

\[
\begin{align*}
\text{Var}(R_t - F_t^T w) &= \frac{1}{T} \sum_{i=1}^{T} (R_t - F_t^T w)^2 - \left( \frac{\sum_{t=1}^{T} (R_t - F_t^T w)}{T} \right)^2 \\
&= \frac{1}{T} \| R - Fw \|^2 - \left( \frac{e^T (R - Fw)}{T} \right)^2 \\
&= \frac{\| R \|^2 - 2R^T Fw + w^T F^T Fw}{T} \\
&\quad - \left( \frac{e^T R)^2 - 2e^T R - w^T F^T e e^T Fw}{T^2} \right)
\end{align*}
\]
Objective function

$$\text{Var}(R_t - F_t^T w) = \frac{\|R\|^2 - 2R^TFw + w^TF^TFw}{T} - \frac{(e^TR)^2 - 2e^TR - w^TF^Te e^TFw}{T^2}$$

Reorganizing the terms as function of $w$: 
Objective function

\[
\text{Var}(R_t - F_t^T w) = \frac{\|R\|^2 - 2 R^T F w + w^T F^T F w}{T} - \frac{(e^T R)^2 - 2 e^T R - w^T F F e^T e F w}{T^2}
\]

Reorganizing the terms as function of \( w \):

\[
\text{Var}(R_t - F_t^T w) = \left( \frac{\|R\|^2}{T} - \frac{(e^T R)^2}{T^2} \right) - 2 \left( \frac{R^T F}{T} - \frac{e^T R}{T^2} e^T F \right) w + w^T \left( \frac{1}{T} F^T F - \frac{1}{T} F F e^T e F \right) w
\]
Objective function

\[
\text{Var}(R_t - F_t^T w) = \frac{\|R\|^2 - 2R^T Fw + w^T F^T Fw}{T} - \frac{(e^T R)^2 - 2e^T R - w^T F^T ee^T Fw}{T^2}
\]

Reorganizing the terms as function of \( w \):

\[
\text{Var}(R_t - F_t^T w) = \left( \frac{\|R\|^2}{T} - \frac{e^T R}{T^2} \right) - 2 \left( \frac{R^T F}{T} - \frac{e^T R}{T^2} e^T F \right) w
+ w^T \left( \frac{1}{T} F^T F - \frac{1}{T} F^T ee^T F \right) w
\]

We have

\[
\frac{1}{T} F^T F - \frac{1}{T} F^T ee^T F = \frac{1}{T} F^T \left( I - \frac{ee^T}{T} \right) F
\]

The matrix \( M = I - ee^T / T \) is symmetric and positive semidefinite (eigenvalues: 0 and 1), and so is \( F^T MF \).

Hence the objective function is convex.