• Book: Optimization Methods in Finance, Gerard Cornuejols and Reha Tütüncü, Carnegie Mellon University, Pittsburgh, USA

• Book: Linear Programming, Vasek Chvatal, McGill University, W.H. Freeman and Company, New York, USA
## Contents

1 **Portfolio Selection and Asset Allocation**
   1.1 Preliminaries and Notation .................................................. 3
   1.2 Mean-variance optimization (MVO) – Markowitz portfolio optimization ........................................ 6
   1.3 Large-Scale Portfolio Optimization ........................................... 11
      1.3.1 Portfolio Optimization with Minimum (and/or Maximum) Transaction Levels .................. 11
      1.3.2 Portfolio Optimization with Transaction Costs ................................................. 12
   1.4 Portfolio Rebalancing Problem .................................................. 13
   1.5 Maximizing the Sharpe Ratio .................................................... 14

2 **Boolean expressions using binary variables** .................................................. 16

3 **Value-at-Risk and Conditional Value-at-Risk** ............................................. 19
   3.1 Value-at-Risk – VaR ............................................................ 19
   3.2 VaR – with continuous and discrete probability distribution ................................................. 22
   3.3 Conditional Value-at-Risk – CVaR ........................................... 22

4 **Optimality Conditions** ................................................................. 26
   4.1 Complementary Slackness ......................................................... 26

5 **Asset Pricing and Arbitrage** .......................................................... 28
   5.1 Pricing and Hedging of Options ................................................ 28
   5.2 Replication Problem ............................................................... 29
   5.3 The Fundamental Theorem of Asset Pricing ..................................... 32

6 **Lagrangian Relaxation** .............................................................. 36

7 **Construction of an Index Fund** ...................................................... 40
1 Portfolio Selection and Asset Allocation

- The theory of optimal selection of portfolios was developed by Harry Markowitz in the 1950s.
- His work formalized the diversification principle in portfolio selection.
- Key assumption: Trades off between expected returns and the perceived risk of portfolios.

1.1 Preliminaries and Notation

- Consider a set of rates of return \( r_t^i \) \((t = 1, \ldots, T, i = 1, \ldots, n)\).
- The value \( r_t^i \) represents the rate of return of a security \( i \) at time \( t \).
- \( T \) represents the temporal horizon under consideration and \( n \) is the number of securities.
- Starting from the time series of the rates of return we can compute:

  arithmetic mean

\[
r_t^a = \frac{1}{T} \sum_{i=1}^{T} r_t^i
\]

  geometric mean (only positive values)

\[
r_t^g = \left( \prod_{i=1}^{T} (1 + r_t^i) \right)^{\frac{1}{T}} - 1
\]

Since the rates of returns are typically not smaller than 1 is it sufficient to add 1 in order to get positive values (then shift back subtracting 1).

Example with 2 rates of returns and a specific security \( i \):

<table>
<thead>
<tr>
<th>( r_t^a )</th>
<th>( r_t^g )</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>0</td>
<td>0.5</td>
</tr>
<tr>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>-0.5</td>
<td>0.0</td>
</tr>
<tr>
<td>0.1</td>
<td>0.25</td>
</tr>
</tbody>
</table>

Table 1: Arithmetic and geometric means

  variance

\[
\sigma_i^2 = \frac{1}{T} \sum_{t=1}^{T} (r_t^i - r_t^a)^2
\]

  standard deviation

\[
\sigma_i = \sqrt{\frac{1}{T} \sum_{t=1}^{T} (r_t^i - r_t^a)^2}
\]
covariance matrix (positive semidefinite)

\[ Q_{ij} = \frac{1}{T} \sum_{t=1}^{T} (r_t^i - r_a^i)(r_t^j - r_a^j) \]

correlation matrix

\[ \rho_{ij} = \frac{Q_{ij}}{\sigma_i \sigma_j} \]

equivalences

\[ Q_{ii} = \sigma_i^2 \]

Example

- Consider the 3 securities: \( S_1, S_2, S_3 \) (\( n = 3 \))
- Consider the price time series of one week (\( T = 5 \))

<table>
<thead>
<tr>
<th></th>
<th>( S_1 )</th>
<th>( S_2 )</th>
<th>( S_3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday</td>
<td>20.2</td>
<td>262.9</td>
<td>100.0</td>
</tr>
<tr>
<td>Tuesday</td>
<td>21.6</td>
<td>268.7</td>
<td>102.3</td>
</tr>
<tr>
<td>Wednesday</td>
<td>20.4</td>
<td>284.0</td>
<td>105.3</td>
</tr>
<tr>
<td>Thursday</td>
<td>21.7</td>
<td>283.1</td>
<td>108.8</td>
</tr>
<tr>
<td>Friday</td>
<td>23.4</td>
<td>282.8</td>
<td>113.0</td>
</tr>
</tbody>
</table>

Table 2: Market Values (Prices)

- Let \( I_t^i \) denote the price of asset \( i = 1, 2, 3 \) at time \( t = 1, \ldots, 5 \), where \( t = 0 \) corresponds to Monday and \( t = 5 \) to Friday (values in Table 2).
- For each asset \( i \), we can compute the rates of returns \( r_t^i \) using the formula:

\[ r_t^i = \frac{I_t^i - I_t^{i-1}}{I_t^{i-1}} \]

- Table 3 reports the rates of returns \( r_t^i \) (\( i = 1, 2, 3 \)) and (\( t = 2, \ldots, 5 \)). Accordingly for this example \( T = 4 \).

<table>
<thead>
<tr>
<th></th>
<th>( S_1 )</th>
<th>( S_2 )</th>
<th>( S_3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tuesday</td>
<td>6.93</td>
<td>2.21</td>
<td>2.30</td>
</tr>
<tr>
<td>Wednesday</td>
<td>-5.56</td>
<td>5.69</td>
<td>2.93</td>
</tr>
<tr>
<td>Thursday</td>
<td>6.37</td>
<td>-0.32</td>
<td>3.32</td>
</tr>
<tr>
<td>Friday</td>
<td>7.83</td>
<td>-0.11</td>
<td>3.86</td>
</tr>
</tbody>
</table>

Table 3: Rates of Return (in percentages): \( r_t^i \)

- From the time series of the rates of return, we can compute the arithmetic means of the rates of return for each asset:

\[ r_a^i = \frac{1}{T} \sum_{t=1}^{T} r_t^i \]
S1    S2    S3

$r_i^a$ (in percentages)  3.90  1.87  3.10

Table 4: Arithmetic means of the rates of return: $r_i^a$

- Table 4 reports the arithmetic means of the rates of return of the example:
- From the time series of the rates of return, we can also compute the geometric means of the rates of return for each asset:

$$r_i^g = \left( \prod_{t=1}^{T} (1 + r_t^i) \right)^{\frac{1}{T}} - 1$$

- Table 5 reports the geometric means of the rates of return of the example:

S1    S2    S3

$r_i^g$ (in percentages)  3.75  1.84  3.10

Table 5: Geometric means of the rates of return: $r_i^g$

- We now compute the covariance matrix:

$$Q_{ij} = \frac{1}{T} \sum_{t=1}^{T} (r_t^i - r^a_i)(r_t^j - r^a_j)$$

<table>
<thead>
<tr>
<th></th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>0.0030</td>
<td>-0.0012</td>
<td>0.0001</td>
</tr>
<tr>
<td>S2</td>
<td>-0.0012</td>
<td>0.0006</td>
<td>-0.0001</td>
</tr>
<tr>
<td>S3</td>
<td>0.0001</td>
<td>-0.0001</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Table 6: Covariance Matrix: $Q_{ij}$

- We can now compute the standard deviation (also called Volatility):

$$\sigma_i = \sqrt{\frac{1}{T} \sum_{t=1}^{T} (r_t^i - r^a_i)^2}$$

<table>
<thead>
<tr>
<th></th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_i$</td>
<td>0.0548</td>
<td>0.0242</td>
<td>0.0057</td>
</tr>
</tbody>
</table>

Table 7: Standard deviation: $\sigma_i$

- We now compute the correlation matrix:

$$\rho_{ij} = \frac{Q_{ij}}{\sigma_i \sigma_j}$$
\[ \begin{array}{ccc}
S_1 & S_2 & S_3 \\
S_1 & 1.0000 & -0.9107 & 0.2166 \\
S_2 & -0.9107 & 1.0000 & -0.5265 \\
S_3 & 0.2166 & -0.5265 & 1.0000 \\
\end{array} \]

Table 8: Correlation Matrix: \( \rho_{ij} \)

1.2 Mean-variance optimization (MVO) – Markowitz portfolio optimization

- Consider now an investor who wants to invest a certain amount of money (budget) in a number of different securities (stocks, bonds, etc.) with random expected returns
- For each security the price time series is available
- For each security \( i = 1, \ldots, n \), we estimates:
  - the expected return \( \mu_i \) (typically the geometric mean of the rates of return \( r_i^g \))
  - the variance \( \sigma_i^2 \) or the standard deviation \( \sigma_i \)
- For any couple of securities \( i \) and \( j \), we estimates:
  - covariance \( Q_{ij} \) or the correlation coefficient \( \rho_{ij} \)
- If we represent the fraction of the total budget invested in security \( i \) by \( x_i \), a portfolio is a vector:
  \[ x = (x_1, \ldots, x_n) \]

expected return of a portfolio

\[ \mathbb{E}[x] = x_1 \mu_1 + \cdots + x_n \mu_n = \mu^\top x \]

where:

\[ \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix} \]

variance of a portfolio

\[ \text{Var}[x] = \sum_{i=1}^{n} \sum_{j=1}^{n} Q_{ij} x_i x_j = \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_{ij} \sigma_i \sigma_j x_i x_j = x^\top Q x \]

where:

\[ Q = \begin{bmatrix} Q_{11} & Q_{12} & \cdots & Q_{1n} \\
Q_{21} & Q_{22} & \cdots & Q_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
Q_{n1} & Q_{n2} & \cdots & Q_{nn} \end{bmatrix} \]

- In linear algebra, a symmetric \( n \cdot n \) real matrix \( M \) is said to be positive semidefinite:
if the scalar $x^T M x$ is non-negative for every non-zero column vector $x$ of $n$ real numbers.

- All the eigenvalues of $M$ are non-negative

The covariance matrix $Q$ is positive semidefinite:

$$
\sum_i \sum_j x_i x_j Q_{ij} = \sum_i \sum_j x_i x_j \frac{1}{T} \sum_{t=1}^T (r^t_i - r^a_i)(r^t_j - r^a_j) =
$$

$$
= \frac{1}{T} \sum_{t=1}^T \sum_i x_i (r^t_i - r^a_i) \sum_j x_j (r^t_j - r^a_j) = \frac{1}{T} \sum_{t=1}^T \left( \sum_j x_j (r^t_j - r^a_j) \right) \left( \sum_i x_i (r^t_i - r^a_i) \right) =
$$

$$
= \frac{1}{T} \sum_{t=1}^T \left( \sum_i x_i (r^t_i - r^a_i) \right) \left( \sum_i x_i (r^t_i - r^a_i) \right) = \frac{1}{T} \sum_{t=1}^T \left( \sum_i x_i (r^t_i - r^a_i) \right)^2 \geq 0
$$

then for any $x$, $x^T Q x \geq 0$.

Portfolio features:

- The portfolio vector $x$ must satisfy $\sum_i x_i = 1$
- There may or may not be additional feasibility constraints
- A feasible portfolio $x$ is called **efficient** if:
  - it has the minimum variance among all portfolios that have at least a certain expected return
  - it has the maximal expected return among all portfolios with less than a certain variance level
- The collection of efficient portfolios forms the **efficient frontier** of the portfolio universe
- The **efficient frontier** is often represented as a curve in a two-dimensional graph where the coordinates of a plotted point corresponds to:
  - the expected return
  - the variance of the efficient portfolio

The MVO problem can be formulated in 2 but equivalent ways:

**First formulation**

- the first formulation results in the problem of finding a minimum variance portfolio of the securities 1 to $n$ that yields at least a target value $R$ of expected return (**target return**).

Convex quadratic programming problem:

$$\min x^T Q x \quad \text{subject to} \quad e^T x = 1, \quad \mu^T x \geq R, \quad x \geq 0$$

where $e$ is an $n$-dimensional vector all of which components are equal to 1

$$e = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$
• The first constraint indicates that the proportions \( x_i \) should sum up to 1
• The second constraint indicates that the expected return is no less than the target value \( R \)
• the objective function corresponds to the total variance of the portfolio
• Nonnegativity constraints on \( x_i \) are introduced to avoid short sales
• Solving this problem for values of \( R \) ranging between \( R_{\text{min}} \) and \( R_{\text{max}} \) one obtains all efficient portfolios
  
  - \( R_{\text{min}} \) can be set to 0 since portfolios with a negative target return can be discarded.
    Attention: if
    \[
    \max_{i=1,...,n} r^g_i < 0
    \]
    the model become infeasible
  
  - \( R_{\text{max}} \) can be set to:
    \[
    \max_{i=1,...,n} r^g_i
    \]
    since in this case the portfolio will consists only of the security \( i^* \) with the largest geometric mean
    of the rates of return. Accordingly the variance of this portfolio will be the \( \sigma^2_{i^*} \).

Second formulation

• the second formulation results in the problem of finding a portfolio of maximum expected return of the
  securities 1 to \( n \) that respects a maximum level of variance \( \bar{R} \).

Convex quadratically constrained programming problem:

\[
\begin{align*}
\max & \mu^T x \\
\text{e}^T x & = 1 \\
\text{x}^T \text{Q} \text{x} & \leq \bar{R} \\
\text{x} & \geq 0
\end{align*}
\]

• this formulation is less used and accordingly in the rest of the section we will focus on the first formulation

Possible additional constraints

• the set of admissible portfolios can be a nonempty polyhedral set represented as

\[
\chi = \{ \text{x} : \text{Ax} = \text{b}, \text{Cx} \geq \text{d} \}
\]

where \( A \) is an \( m \cdot n \) matrix, \( b \) is an \( m \)-dimensional vector, \( C \) is a \( p \cdot n \) matrix and \( d \) is a \( p \)-dimensional
vector.

• In this case it is necessary to add the following constraints to the models

\[
\begin{align*}
\text{Ax} & = \text{b} \\
\text{Cx} & \geq \text{b}
\end{align*}
\]

• For example, if regulations or investor preferences limit the amount of investment in a subset of the
  securities, the model can be augmented with a linear constraint to reflect such a limit.

• In principle, any linear constraint can be added to the model without making it significantly harder to
  solve.
Markowitz MVO model example

- Consider the 3 securities: $S_1, S_2, S_3, S_4$ ($n = 4$)
- Consider the following price time series of 4 days ($T = 4$)

<table>
<thead>
<tr>
<th></th>
<th>$S_1$</th>
<th>$S_2$</th>
<th>$S_3$</th>
<th>$S_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday</td>
<td>100.0</td>
<td>100.0</td>
<td>1.0</td>
<td>2.0</td>
</tr>
<tr>
<td>Tuesday</td>
<td>110.0</td>
<td>90.0</td>
<td>1.1</td>
<td>1.5</td>
</tr>
<tr>
<td>Wednesday</td>
<td>132.0</td>
<td>99.0</td>
<td>1.1</td>
<td>1.5</td>
</tr>
<tr>
<td>Thursday</td>
<td>145.2</td>
<td>108.9</td>
<td>1.1</td>
<td>2.0</td>
</tr>
</tbody>
</table>

- The rates of returns $r_t^i$ ($i = 1, 2, 3, 4$) and ($t = 2, 3, 4$) are: (this example $T = 3$).

<table>
<thead>
<tr>
<th></th>
<th>$S_1$</th>
<th>$S_2$</th>
<th>$S_3$</th>
<th>$S_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tuesday</td>
<td>10.00</td>
<td>-10.00</td>
<td>10.00</td>
<td>-25.00</td>
</tr>
<tr>
<td>Wednesday</td>
<td>20.00</td>
<td>10.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Thursday</td>
<td>10.00</td>
<td>10.00</td>
<td>0.00</td>
<td>33.33</td>
</tr>
</tbody>
</table>

- The arithmetic means of the rates of return of the example are:

<table>
<thead>
<tr>
<th></th>
<th>$S_1$</th>
<th>$S_2$</th>
<th>$S_3$</th>
<th>$S_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_a^i$ (in percentages)</td>
<td>13.33</td>
<td>3.33</td>
<td>3.33</td>
<td>2.78</td>
</tr>
</tbody>
</table>

- The geometric means of the rates of return of the example are:

<table>
<thead>
<tr>
<th></th>
<th>$S_1$</th>
<th>$S_2$</th>
<th>$S_3$</th>
<th>$S_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_g^i$ (in percentages)</td>
<td>13.24</td>
<td>2.88</td>
<td>3.23</td>
<td>0.00</td>
</tr>
</tbody>
</table>

- We now compute the covariance matrix:

<table>
<thead>
<tr>
<th></th>
<th>$S_1$</th>
<th>$S_2$</th>
<th>$S_3$</th>
<th>$S_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_1$</td>
<td>0.0022</td>
<td>0.0022</td>
<td>-0.0011</td>
<td>-0.0009</td>
</tr>
<tr>
<td>$S_2$</td>
<td>0.0022</td>
<td>0.0089</td>
<td>-0.0044</td>
<td>0.0185</td>
</tr>
<tr>
<td>$S_3$</td>
<td>-0.0011</td>
<td>-0.0044</td>
<td>0.0022</td>
<td>-0.0093</td>
</tr>
<tr>
<td>$S_4$</td>
<td>-0.0009</td>
<td>0.0185</td>
<td>-0.0093</td>
<td>0.0571</td>
</tr>
</tbody>
</table>

- We construct now the Markowitz’s MVO model to construct a portfolio of four securities: $S_1, S_2, S_3, S_4$ of the previous example.

- $R_{min}$ can be set to: 0.0

- $R_{max}$ can be set to: 0.1324, i.e., $\mu_1$
• For an expected target return $R$ the QP model reads as follows:

$$\begin{align*}
\min & \quad \left[0.0044 \cdot x_1^2 + 0.0088 \cdot x_1 \cdot x_2 + 0.0044 \cdot x_1 \cdot x_3 - 0.0037 \cdot x_1 \cdot x_4 + 0.0177 \cdot x_2^2 - 0.0177 \cdot x_2 \cdot x_3 + 0.0740 \cdot x_2 \cdot x_4 + 0.0044 \cdot x_3^2 + 0.0370 \cdot x_3 \cdot x_4 + 0.1141 \cdot x_4^2\right] / 2 \\
& \quad x_1 + x_2 + x_3 + x_4 = 1 \\
& \quad 0.1324 \cdot x_1 + 0.0288 \cdot x_2 + 0.0323 \cdot x_3 + 0.0 \cdot x_4 \geq R \\
& \quad x_1, x_2, x_3, x_4 \geq 0
\end{align*}$$

• Solving the model for $R = 0.06$ to $R = 0.01324$ with increments of 0.01 we get the following optimal portfolios and the corresponding variance.

<table>
<thead>
<tr>
<th>Target Return $R$</th>
<th>Variance</th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
<th>$x_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.06</td>
<td>0.00000</td>
<td>0.3279</td>
<td>0.0000</td>
<td>0.5738</td>
<td>0.0984</td>
</tr>
<tr>
<td>0.07</td>
<td>0.00003</td>
<td>0.4044</td>
<td>0.0000</td>
<td>0.5102</td>
<td>0.0854</td>
</tr>
<tr>
<td>0.08</td>
<td>0.00015</td>
<td>0.4991</td>
<td>0.0000</td>
<td>0.4315</td>
<td>0.0694</td>
</tr>
<tr>
<td>0.09</td>
<td>0.00035</td>
<td>0.5939</td>
<td>0.0000</td>
<td>0.3528</td>
<td>0.0533</td>
</tr>
<tr>
<td>0.10</td>
<td>0.00064</td>
<td>0.6886</td>
<td>0.0000</td>
<td>0.2741</td>
<td>0.0373</td>
</tr>
<tr>
<td>0.11</td>
<td>0.00103</td>
<td>0.7833</td>
<td>0.0000</td>
<td>0.1954</td>
<td>0.0212</td>
</tr>
<tr>
<td>0.12</td>
<td>0.00150</td>
<td>0.8781</td>
<td>0.0000</td>
<td>0.1167</td>
<td>0.0052</td>
</tr>
<tr>
<td>0.13</td>
<td>0.00207</td>
<td>0.9763</td>
<td>0.0000</td>
<td>0.0237</td>
<td>0.0000</td>
</tr>
<tr>
<td>0.1324</td>
<td>0.00222</td>
<td>1.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Table 9: Efficient Portfolios

• Efficient frontier

![Efficient frontier](image-url)
Asset allocation

- Asset allocation problems have the same mathematical structure as portfolio selection problems.
- In these problems the objective is not to choose a portfolio of stocks (or other securities) but to determine the optimal investment among a set of asset classes.
- Examples of asset classes are large capitalization stocks, small capitalization stocks, foreign stocks, government bonds, corporate bonds, etc.
- After estimating the expected returns, variances, and covariances for different asset classes, one can formulate a Markowitz model and obtain efficient portfolios of these asset classes.

1.3 Large-Scale Portfolio Optimization

- We consider now the practical issues that arise when the Mean-Variance model is used to construct a portfolio from a large underlying family of assets or stocks.
- the number of assets or stocks $n$ may be in the hundreds or thousands.
- Diversification – In general, there is no reason to expect that solutions to the Markowitz model will be well diversified portfolios.
- Practitioners often use additional constraints to ensure that the chosen portfolio is well diversified.

1.3.1 Portfolio Optimization with Minimum (and/or Maximum) Transaction Levels

When solving the classical Markowitz model, the optimal portfolio often contains $x_i$ that are too small or too big. In practice, one would like a solution of with the following properties.

With the property that:

$$x_i \leq l_{i_{max}} (A)$$

where $l_{i_{max}}$ are given maximum transaction levels.

- this constraint alone can be easily considered as follow

$$x_i \leq l_{i_{max}} \quad i = 1, \ldots, n$$
With the property that:

\[ \text{if } x_i > 0 \Rightarrow x_i \geq l_{i \text{ min}}^i \ (B) \]

where \( l_{i \text{ min}}^i \) are given minimum transaction levels.

- This constraint states that, if an investment is made in a stock, then it must be large enough, for example at least 100 shares.
- Because this constraint is not a simple linear constraint (logic constraint), it cannot be handled directly by a solver.

**BQP formulation**

- The Mean-Variance (MV) portfolio problem with minimum and maximum transaction level requires optimally allocating wealth among a set \( N \) of assets in order to obtain a prescribed level of target return \( R \) while minimizing the risk as measured by the variance of the portfolio.

\[
\begin{align*}
\min & \sum_{i=1}^{n} \sum_{j=1}^{n} Q_{ij}x_i x_j \\
\sum_{i=1}^{n} \mu_i x_i & \geq R \\
\sum_{i=1}^{n} x_i & = 1 \\
x_i & \geq l_{i \text{ min}}^i u_i \quad i = 1, \ldots, n \\
x_i & \leq l_{i \text{ max}}^i u_i \quad i = 1, \ldots, n \\
x_i & \geq 0 \quad i = 1, \ldots, n \\
u_i & \in \{0, 1\} \quad i = 1, \ldots, n
\end{align*}
\]

- where \( \mu_i, l_{i \text{ min}}^i \) and \( l_{i \text{ max}}^i \) are respectively the expected unitary return and the minimum and maximum transaction level for asset \( i \),
- This apparently simple model is rather demanding for general-purpose (MIQP) solvers, since the root node gaps of the ordinary continuous relaxation are huge.
- In addition classical polyhedral approaches are not effective to improve the lower bounds.

**1.3.2 Portfolio Optimization with Transaction Costs**

- We can add a portfolio turnover constraint to ensure that the change between the current holdings \( x_0 \) and the desired portfolio \( x \) is bounded by \( h \)
- This constraint is essential when solving large mean-variance models since the covariance matrix (A square matrix which does not have an inverse. A matrix is singular if and only if its determinant is zero.) is almost singular in most practical applications and hence the optimal decision can change significantly with small changes in the problem data.
- To avoid big changes when reoptimizing the portfolio, turnover constraints can be imposed.
- Let \( x^0 \) denote the current portfolio \( (x^0_i \ i = 1, \ldots, n) \)
Let $y_i$ be the amount of asset $i$ bought and $z_i$ the amount sold, and $h$ a bound on the total amount bought and sold:

\begin{align*}
  x_i - x_i^0 & \leq y_i, \quad y_i \geq 0, \quad (27) \\
  x_i^0 - x_i & \leq z_i, \quad z_i \geq 0, \quad (28) \\
  \sum_{i=1}^n y_i + z_i & \leq h \quad (29)
\end{align*}

We can introduce transaction costs directly into the model

Suppose that there is a transaction cost $t_i$ proportional to the amount of asset $i$ bought, and a transaction cost $t'_i$ proportional to the amount of asset $i$ sold

Then a re-optimized portfolio is obtained by solving

\begin{align*}
  \min & \sum_{i=1}^n \sum_{j=1}^n Q_{ij} x_i x_j \\
  \sum_{i=1}^n \mu_i x_i - t_i y_i - t'_i z_i & \geq R \quad (31) \\
  \sum_{i=1}^n x_i & = 1 \quad (32) \\
  x_i - x_i^0 & \leq y_i \quad i = 1, \ldots, n \quad (33) \\
  x_i^0 - x_i & \leq z_i \quad i = 1, \ldots, n \quad (34) \\
  \sum_{i=1}^n y_i + z_i & \leq h \quad (35) \\
  x_i & \geq 0 \quad i = 1, \ldots, n \quad (36) \\
  y_i & \geq 0 \quad i = 1, \ldots, n \quad (37) \\
  z_i & \geq 0 \quad i = 1, \ldots, n \quad (38)
\end{align*}

1.4 Portfolio Rebalancing Problem

we consider an approach which tries to rebalance the portfolio at minimum cost

This approach assumes that we have identified important characteristics of the market index we would like to achieve. Such characteristics might be:

- the fraction $f_k$ of companies with market capitalization in various range
- the fraction $f_k$ of companies that pay no dividends
- ...

Let us assume that there are $m$ such characteristics that we would like our index fund to track as well as possible

Let $a_{ki} = 1$ if company $i$ has characteristic $k$ and $0$ if it does not

Let $x_i$ denote the optimum weight of asset $i$ in the portfolio. Assume that initially, the portfolio has weights $x_i^0$. 

Let $y_i$ denote the fraction of asset $i$ bought and $z_i$ the fraction sold. The problem of rebalancing the portfolio at minimum cost is the following:

$$\min \sum_{i=1}^{n} (y_i + z_i) \tag{39}$$

$$\sum_{i=1}^{n} a_{ki} x_i = f_i \quad k = 1, \ldots, m \tag{40}$$

$$\sum_{i=1}^{n} x_i = 1 \tag{41}$$

$$x_i - x_i^0 \leq y_i \quad i = 1, \ldots, n \tag{42}$$

$$x_i^0 - x_i \leq z_i \quad i = 1, \ldots, n \tag{43}$$

$$x_i \geq 0 \quad i = 1, \ldots, n \tag{44}$$

$$y_i \geq 0 \quad i = 1, \ldots, n \tag{45}$$

$$z_i \geq 0 \quad i = 1, \ldots, n \tag{46}$$

1.5 Maximizing the Sharpe Ratio

- It is a method to uniquely define the “optimal” portfolio
- Recall that we denote with $R_{min}$ and $R_{max}$ the minimum and maximum target expected returns of efficient portfolios
- Let us define the function $\sigma(R)$ where $R$ is the target expected return:

$$\sigma(R) : [R_{min}, R_{max}] \to \mathbb{R}$$

$$\sigma(R) := (x_R^T Q x_R)$$

$\Rightarrow x_R$ denotes the unique optimal solution of the MVO problem for each $R$

- Since we assumed that $Q$ is positive definite, it is possible to show that the function $\sigma(R)$ is strictly convex in its domain
- the efficient frontier is the graph

$$E = \{(R, \sigma(R)) : R \in [R_{min}, R_{max}]\}$$

- A set of “optimal” portfolios – Which one is the “best”?
- We now consider a riskless asset whose expected return is $r_f \geq 0$. We can assume that $r_f < R_{min}$, since the portfolio $x_{min}$ has a positive risk associated with it while the riskless asset does not.
- The Sharpe ratio measures the excess return (or risk premium) per unit of variance (risk) of a portfolio (named after William F. Sharpe)

$$h(x) = \frac{\mu^T x - r_f}{x^T Q x}$$

The portfolio that maximizes the Sharpe ratio is found by solving the following problem:

$$\max \frac{\mu^T x - r_f}{x^T Q x} \tag{47}$$

$$e^T x = 1 \tag{48}$$

$$x \geq 0 \tag{49}$$
• \( x^\top Q x > 0 \)

since \( Q \) is a positive definite matrix, i.e. \( x^\top Q x > 0 \) for every non-zero column vector \( x \) of \( n \) real numbers. Or equivalently all the eigenvalues of \( Q \) are positive.

• In this form, this problem is not easy to solve. Although it has a nice polyhedral feasible region, its objective function is somewhat complicated, and worse, is possibly non-concave. It is not a convex optimization problem.

• An effective way of computing the optimal risky portfolio is based instead on the following graphic interpretation of the Sharpe Ratio

**Graphic interpretation of the Sharpe Ratio**

• Return/risk profiles of different combinations of a risky portfolio with the riskless asset can be represented as a straight line – capital allocation line (CAL) –. This line start in \( r_f \) (all budget allocated in the riskless asset) and it passes thought \( x^\top R Q x \) (all budget allocated in the portfolio)

• The optimal CAL never goes above a point on the efficient frontier (otherwise you can have the same return with a smaller risk). The optimal CAL touches in only one point the efficient frontier.

• The point where the optimal CAL touches the efficient frontier corresponds to the optimal risky portfolio.

<table>
<thead>
<tr>
<th><img src="image" alt="Capital Allocation Line" /></th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 3: Capital Allocation Line</td>
</tr>
</tbody>
</table>

• Alternatively, one can think of the optimal CAL as the CAL with the biggest slope \( \left( \frac{df(x)}{dx} \right) \). Mathematically, this can be expressed as the portfolio \( x \) that maximizes the quantity

\[
h(x) = \frac{\mu^\top x - r_f}{x^\top Q x}
\]

• The effective strategy to find the portfolio maximizing the Sharpe ratio (optimal risky portfolio) is the following:
  
  – First, one traces out the efficient frontier on a two dimensional target return vs. variance graph
  – Then, the point on this graph corresponding to the optimal risky portfolio is found as the tangency point of the line going through the point representing the riskless asset and tangent to the efficient frontier.
  – Once this point is identified, one can recover the composition of this portfolio from the information generated and recorded while constructing the efficient frontier.
2 Boolean expressions using binary variables

- A Boolean expression, also called propositional logic formula, is built from

1) boolean-valued functions (also called proposition)
   - functions \( f : X \rightarrow B \)
   - \( X \) is an arbitrary set
   - \( B \) is a boolean domain \( B \in \{ \text{TRUE}, \text{FALSE} \} \)
   - example: given a set \( N \) of items, a function that returns \( \text{TRUE} \) if an item is selected and \( \text{FALSE} \) otherwise,
     \[ X \equiv N \quad \text{and} \quad B \in \{ \text{TRUE}, \text{FALSE} \} \]
   - a boolean-valued functions can be represented by a binary variable:
     \[ x \in \{ 0, 1 \}, \quad \text{if TRUE} \rightarrow x = 1, \quad \text{else} (\text{if FALSE}) \rightarrow x = 0 \]

2) boolean operators:
   - \( \text{AND} \) – conjunction – denoted also by \( \land \)
   - \( \text{OR} \) – disjunction – denoted also by \( \lor \)
   - the negation of a boolean-valued functions represented by a binary variable is \( \text{NOT} f \)
   - it corresponds to the binary variable \( 1 - x \)

3) parentheses

- A Boolean expression is said to be satisfiable if it can be made \( \text{TRUE} \) by assigning appropriate logical values (i.e., \( \text{TRUE}, \text{FALSE} \)) to its binary variables

- Example of Boolean expression:
  \[(x_a \text{ OR } x_b) \text{ AND } (x_c \text{ OR } x_d) \text{ AND } (\text{ NOT } x_e)\]

- Conjunction:
  \[x_a \text{ AND } x_b \Leftrightarrow \begin{cases} x_a \geq 1 \\ x_b \geq 1 \end{cases}\]

- Disjunction:
  \[x_a \text{ OR } x_b \Leftrightarrow x_a + x_b \geq 1\]

- Remarque : the boolean operator \( \text{AND} \) come first than \( \text{OR} \). \( x_a \text{ OR } x_b \text{ AND } x_c \) is equivalent to \( x_a \text{ OR } (x_b \text{ AND } x_c) \)
Conjunctive Normal Form (CNF)

- A Boolean expression is said to in Conjunctive Normal Form if it is expressed as conjunctions of disjunctions
- examples:
  - $x_a \text{ AND } x_b$
  - $(x_a \text{ OR } x_b) \text{ AND } x_c$
- Equivalences to get CNFs:
  - $\neg (\neg x_a) = x_a$
  - $\neg (x_a \text{ OR } x_b) = (\neg x_a) \text{ AND } (\neg x_b)$
  - $\neg (x_a \text{ AND } x_b) = (\neg x_a) \text{ OR } (\neg x_b)$
  - $(x_a \text{ AND } x_b) = (x_a \text{ AND } x_b) \text{ OR } (x_a \text{ AND } x_c)$
  - $x_a \text{ OR } (x_b \text{ AND } x_c) = (x_a \text{ OR } x_b) \text{ AND } (x_a \text{ OR } x_c)$

Transforming CNF into linear inequalities

- for each disjunction we have an inequality constraint $\geq 1$ (the most simple CNF is composed by only one disjunction)
- the left hand side of the inequality is written replacing:
  - OR with +
  - each (NOT $x_a$) with $(1 - x_a)$

Logical Implications between binary variables

- the logical implication $x_a \Rightarrow x_b$ is expressed by the following boolean expression $(\neg x_a) \text{ OR } x_b$
- this can be written as a linear constraints using the binary variables as follow:
  $$(1 - x_a) + x_b \geq 1 \quad (x_b \geq x_a)$$
- Attention that $x_a \Rightarrow x_b$ implies also $\neg x_b \Rightarrow \neg x_a = (\neg (\neg x_b)) \text{ OR } \neg x_a = (\neg x_a) \text{ OR } x_b$

examples

1. $(x_a \text{ AND } x_b) \Rightarrow x_c$

   $$\neg (x_a \text{ AND } x_b) \text{ OR } x_c$$
   $$(\neg x_a) \text{ OR } (\neg x_b) \text{ OR } x_c$$
   $$(1 - x_a) + (1 - x_b) + x_c \geq 1$$
   $$x_a + x_b - x_c \leq 1$$

2. $(x_a \text{ OR } x_b) \Rightarrow x_c$

   $$\neg (x_a \text{ OR } x_b) \text{ OR } x_c$$
   $$((\neg x_a) \text{ OR } (\neg x_b)) \text{ OR } x_c$$
   $$((\neg x_a) \text{ OR } (x_c)) \text{ AND } ((\neg x_b) \text{ OR } (x_c))$$
   $$\begin{cases} 1 - x_a + x_c \geq 1 \\ 1 - x_b + x_c \geq 1 \end{cases}$$
(3) $x_a \Rightarrow (x_b \text{ AND } x_c)$

\[
\begin{align*}
\text{NOT } (x_a) & \text{ OR } (x_b \text{ AND } x_c) \\
((\text{ NOT } x_a) \text{ OR } (x_b)) & \text{ AND } ((\text{ NOT } x_a) \text{ OR } (x_c)) \\
\begin{cases}
1 - x_a + x_b & \geq 1 \\
1 - x_a + x_c & \geq 1
\end{cases}
\end{align*}
\]

(4) $x_a \Rightarrow (x_b \text{ OR } x_c)$

\[
\begin{align*}
\text{NOT } (x_a) & \text{ OR } (x_b \text{ OR } x_c) \\
1 - x_a + x_b + x_c & \geq 1 \\
(x_b + x_c & \geq x_a)
\end{align*}
\]

(5) $x_a \text{ AND } (\text{ NOT } x_b) \text{ AND } x_c \Rightarrow \text{ NOT } x_d$

\[
\begin{align*}
\text{NOT } (x_a \text{ AND } (\text{ NOT } x_b) \text{ AND } x_c) & \text{ OR } (\text{ NOT } x_d) \\
(\text{ NOT } x_a) & \text{ OR } x_b \text{ OR } (\text{ NOT } x_c) \text{ OR } (\text{ NOT } x_d) \\
1 - x_a + x_b + 1 - x_c + 1 - x_d & \geq 1 \\
(x_a - x_b + x_c + x_d & \leq 2)
\end{align*}
\]
3 Value-at-Risk and Conditional Value-at-Risk

- Financial activities involve risk
- Financial and other institutions must manage risk using sophisticated mathematical techniques
- Necessity of quantitative risk measures that adequately reflect the vulnerabilities of a company
- Value-at-Risk and Conditional Value-at-Risk are widely used measure of risk in finance
- They can be used instead of the variance of a portfolio as in the Markowitz model
- This is achieved through stochastic programming
- The random events are modeled by a large but finite set of scenarios, leading to a linear programming equivalent of the original stochastic program.

3.1 Value-at-Risk – VaR

**Definition** – Take a random variable $X$ that represents for example the loss of an investment over a fixed period of time (a negative value for $X$ indicates gains). Given a probability level $\alpha$, $\alpha$-VaR of the random variable $X$ is given by the following relation:

$$ VaR_\alpha(X) = \min \{ \gamma : P(X \leq \gamma) \geq \alpha \} $$

**VaR**: represents the predicted maximum loss with a specified probability level (e.g., 95%) over a certain period of time (e.g., one day).

![Figure 4: 0.95-VaR on a portfolio loss distribution plot](image)

Figure 4: 0.95-VaR on a portfolio loss distribution plot
Example of application

- take a generic portfolio
- you know its current market value, at the beginning of the day while it is not known its market value at the end of the day
- The investment bank that holds this portfolio may declare that its portfolio has a 1-day VaR of 0.1 million with a confidence level of 95%
- It means that the bank expects that, with a probability of 95%, the portfolio’s loss will be less than 0.1 million during the day
  - This implies that the bank expects that the loss of its portfolio at the end of the day will be more than 0.1 million, with a probability of 5%
  - So the bank expects that 5 times out of 100, the portfolio loss will be greater than 0.1 million, while below this threshold will be 95 times out of 100

Features:

- VaR is widely used by people in the financial industry and VaR calculators are common features in most financial software
- VaR can be estimated either parametrically (for example, variance-covariance VaR or delta-gamma VaR) or nonparametrically (for examples, historical simulation VaR or resampled VaR)
- 85% of large banks use historical simulation. The other 15% used Monte Carlo methods

Problems:

- it lacks of sub-additivity
- a sub-additive function $f(x_1 + x_2) \leq f(x_1) + f(x_2) \forall x_1, x_2$
- a good Risk measures should instead respect the principle that diversification reduces risk
- When VaR is computed by generating scenarios, it turns out to be a non-smooth and non-convex function.
- Therefore, when one tries to optimize VaR computed in this manner, the optimization is difficult (global optimization algorithms)
- Another criticism on VaR is that it pays no attention to the magnitude of losses beyond the VaR value.

Example:

- Consider two independent and identical investment opportunities:
  - 1 of gain with probability 96%
  - 2 of loss with probability 4%
- Then, 0.95-VaR for both investments are -1.
- Now consider the sum of these two investment opportunities. Because of independence, this sum has the following loss distribution:
  - 4 of loss with probability 1.6% (0.04 * 0.04)
  - 1 of loss with probability 7.68% (2 * 0.96 * 0.04)
  - 2 of gain with probability 92.16% (0.96 * 0.96)
- Therefore, the 0.95-VaR of the sum of the two investments is 1, which exceeds −2, the sum of the 0.95-VaR values for individual investments (lacks of sub-additivity).
Example of calculation

- consider the following losses of a portfolio over a time window of 20 days

<table>
<thead>
<tr>
<th>day</th>
<th>loss</th>
<th>day</th>
<th>loss</th>
<th>day</th>
<th>loss</th>
<th>day</th>
<th>loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>6</td>
<td>-30</td>
<td>11</td>
<td>20</td>
<td>16</td>
<td>-5</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>7</td>
<td>-40</td>
<td>12</td>
<td>15</td>
<td>17</td>
<td>-10</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>8</td>
<td>10</td>
<td>13</td>
<td>-5</td>
<td>18</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>15</td>
<td>9</td>
<td>15</td>
<td>14</td>
<td>-10</td>
<td>19</td>
<td>15</td>
</tr>
<tr>
<td>5</td>
<td>-30</td>
<td>10</td>
<td>10</td>
<td>15</td>
<td>-10</td>
<td>20</td>
<td>10</td>
</tr>
</tbody>
</table>

- we can now compute and plot the portfolio loss distribution and cumulative loss distribution

<table>
<thead>
<tr>
<th>loss</th>
<th>probability</th>
<th>cumulative</th>
</tr>
</thead>
<tbody>
<tr>
<td>-40</td>
<td>5.00</td>
<td>5.00</td>
</tr>
<tr>
<td>-30</td>
<td>10.00</td>
<td>15.00</td>
</tr>
<tr>
<td>-10</td>
<td>15.00</td>
<td>30.00</td>
</tr>
<tr>
<td>-5</td>
<td>10.00</td>
<td>40.00</td>
</tr>
<tr>
<td>10</td>
<td>30.00</td>
<td>70.00</td>
</tr>
<tr>
<td>15</td>
<td>20.00</td>
<td>90.00</td>
</tr>
<tr>
<td>20</td>
<td>10.00</td>
<td>100.00</td>
</tr>
</tbody>
</table>

portfolio loss distribution

portfolio cumulative loss distribution
• the 0.95-VaR is the smallest loss such that the cumulative probability is greater or equal to 95, 0.95-VaR = 20
• the 0.70-VaR is the smallest loss such that the cumulative probability is greater or equal to 70, 0.70-VaR = 10

3.2 VaR – with continuous and discrete probability distribution

• We consider a portfolio of assets with random returns
• We denote the portfolio choice vector with \( x \) and the random events by the vector \( y \)
• Let \( f(x, y) \) denote the loss function when we choose the portfolio \( x \) from a set \( X \) of feasible portfolios and \( y \) is the realization of the random events
• We assume that the random vector \( y \) has a probability density function denoted by \( p(y) \)

For a fixed decision vector \( x \), we compute the cumulative distribution function of the loss associated with that vector \( x \):

Continuous probability distribution:

\[
\Psi(x, y) = \int_{f(x,y) < \gamma} p(y) \, dy
\]

Discrete probability distributions:

\[
\Psi(x, y) = \sum_{j: f(x,y_j) < \gamma} p_j
\]

Then, for a given confidence level \( \alpha \), the \( \alpha \)-VaR associated with portfolio \( x \) is given as:

\[
VaR_\alpha(x) = \min\{ \gamma \in R : \Psi(x, y) \geq \alpha \}
\]

3.3 Conditional Value-at-Risk – CVaR

We define the \( \alpha \)-CVaR associated with portfolio \( x \) as:

\[
CVaR_\alpha(x) = \frac{1}{1 - \alpha} \int_{f(x,y) \geq VaR_\alpha(x)} f(x,y) p(y) \, dy
\]

For a discrete probability distribution (where event \( y_j \) occurs with probability \( p_j \), for \( j = 1, \ldots, n \)), the above definition of CVaR becomes:
CVaR, \( \alpha \) = \frac{1}{1 - \alpha} \sum_{j : f(x, y_j) \geq \text{VaR}_\alpha(x)} p_j f(x, y_j)

Features:

- CVaR of a portfolio is always at least as big as its VaR
- Consequently, portfolios with small CVaR also have small VaR
- However, in general minimizing CVaR and VaR are not equivalent

Example:
- Suppose we are given the loss function \( f(x, y) \) for a given decision \( x \) as:
  - \( f(x, y) = -y_j \) where \( y_j = 75 - j \) with probability 1% for \( j = 0, \ldots, 99 \)
  - We would like to determine the maximum loss incurred with 95% probability
  - This is the Value-at-Risk \( \text{VaR}_\alpha(x) \) for \( \alpha = 95\% \)
  - We have \( \text{VaR}_{95\%}(x) = 19 \) since the loss is less or equal to 19 with probability 95%
  - To compute the Conditional Value-at-Risk, we use the above formula:
    - \( \text{CVaR}_{95\%}(x) = \frac{1}{0.05} (19 + 20 + 21 + 22 + 23 + 24) \times 1\% = 25.8 \)

portfolio loss function
Since the definition of CVaR involves the VaR function explicitly, it is difficult to work with and optimize this function. Instead, we consider the following simpler auxiliary function:

$$ F_\alpha(x, \gamma) = \gamma + \frac{1}{1 - \alpha} \int_{f(x,y) \geq \gamma} (f(x,y) - \gamma) p(y) \, dy $$

Alternatively, we can write

$$ F_\alpha(x, \gamma) = \gamma + \frac{1}{1 - \alpha} \int (f(x,y) - \gamma)^+ p(y) \, dy $$

This function, viewed as a function of $\gamma$, has the following important properties that makes it useful for the computation of VaR and CVaR:

- $F_\alpha(x, \gamma)$ is a convex function of $\gamma$
- $\text{VaR}_\alpha (x)$ is a minimizer over $\gamma$ of $F_\alpha(x, \gamma)$
- The minimum value over $\gamma$ of the function $F_\alpha(x, \gamma)$ is $\text{CVaR}_\alpha (x)$
As a consequence of the listed properties, we immediately deduce that, in order to minimize \( \text{CVaR}_\alpha(x) \) over \( x \), we need to minimize the function \( F_\alpha(x, \gamma) \) with respect to \( x \) and \( \gamma \) simultaneously

\[
\min_{x \in X} \text{CVaR}_\alpha(x) = \min_{x \in X, \gamma} F_\alpha(x, \gamma)
\]

- Consequently, we can optimize CVaR directly, without needing to compute VaR first
- If the loss function \( f(x, y) \) is a convex (linear) function of the portfolio variables \( x \), then \( F_\alpha(x, \gamma) \) is also a convex (linear) function of \( x \)
- In this case, provided the feasible portfolio set \( X \) is also convex, the optimization problems above are smooth convex optimization problems that can be solved using well known optimization techniques for such problems.

**Scenarios**

- Often it is not possible or desirable to compute/determine the joint density function \( p(y) \) of the random events in our formulation
- Instead, we may have a number of scenarios (with equal probability), say \( y_s \) for \( s = 1, \ldots, S \), which may represent some historical values of the random events or some values obtained via computer simulation.

In this case, we obtain the following approximation to the function:

\[
\tilde{F}_\alpha(x, \gamma) = \gamma + \frac{1}{(1 - \alpha)S} \sum_{s=1}^{S} (f(x, y_s) - \gamma)
\]

Now, the problem:

\[
\min_{x \in X} \text{CVaR}_\alpha(x) = \min_{x \in X, \gamma} \gamma + \frac{1}{(1 - \alpha)S} \sum_{s=1}^{S} (f(x, y_s) - \gamma)
\]

To solve this optimization problem, we introduce artificial variables \( z_s \) to replace \( (f(x, y_s) - \gamma) \), then obtaining:

\[
\min_{x, z, \gamma} \gamma + \frac{1}{(1 - \alpha)S} \sum_{s=1}^{S} z_s \quad (50)
\]

\[
z_s \geq 0 \quad s = 1, \ldots, S \quad (51)
\]

\[
z_s \geq f(x, y_s) - \gamma \quad s = 1, \ldots, S \quad (52)
\]

\[
x \in X \quad (53)
\]

In the case that \( f(x, y) \) is linear in \( x \), all the constraints are linear constraints and therefore the problem is a linear programming problem that can be solved using the simplex method or alternative LP algorithms.
4 Optimality Conditions

4.1 Complementary Slackness

Theorem 4.1. A feasible solution \((x_1^*, \ldots, x_n^*)\) of the Primal Problem \(P\) is optimal if and only if there are numbers \((y_1^*, \ldots, y_m^*)\) such that

\[
\sum_{i=1}^{m} a_{ij} y_i^* = c_j \quad \text{whenever} \quad x_j^* > 0
\]

\[
y_i^* = 0 \quad \text{whenever} \quad \sum_{j=1}^{n} a_{ij} x_j^* < b_i
\]

and such that

\[
\sum_{i=1}^{m} a_{ij} y_i^* \geq c_j \quad \text{for all} \quad j = 1, \ldots, n
\]

\[
y_i^* \geq 0 \quad \text{for all} \quad i = 1, \ldots, m.
\]

Example with \(n = 2\) and \(m = 2\):

\[
\text{Z}(LP) = \max \begin{cases} 
    x_1 + \frac{164}{100} x_2 \\
    50x_1 + 31x_2 \leq 250 \\
    -3x_1 + 2x_2 \leq 4 \\
    x_1, x_2 \geq 0
\end{cases}
\]

\[
A = \begin{bmatrix} 50 & 31 \\ -3 & 2 \end{bmatrix}, \quad b = \begin{bmatrix} 250 \\ 4 \end{bmatrix}, \quad c = \begin{bmatrix} \frac{164}{100} \\ \frac{100}{95} \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}
\]

- Feasible solution: \(x^* = \begin{bmatrix} 376 \\ 193 \\ 950 \\ 193 \end{bmatrix}\) optimal?

- \(y_1^* = 0\) if \(\sum_{j=1}^{n} a_{ij} x_j^* < b_i\)

\[
\begin{cases} 
    50 \cdot \frac{376}{193} + 31 \cdot \frac{950}{193} = 250 \\
    -3 \cdot \frac{376}{193} + 2 \cdot \frac{950}{193} = 4
\end{cases} \quad \rightarrow \text{no y variables set to 0}
\]

- \(\sum_{i=1}^{m} a_{ij} y_i^* = c_j\) if \(x_j^* > 0\)

\[
\begin{cases} 
    50y_1 - 3y_2 = \frac{1}{100} \\
    31y_1 + 2y_2 = \frac{64}{100}
\end{cases} \quad y^* = \begin{bmatrix} \frac{98}{193} \\ \frac{1}{193} \end{bmatrix} \rightarrow \text{Optimal}
\]

- \(y_i^* \geq 0\) forall \(i = 1, \ldots, m \rightarrow \text{OK}\)

- \(\sum_{i=1}^{m} a_{ij} y_i^* \geq c_j\)forall \(j = 1, \ldots, n\)

\[
\begin{cases} 
    50 \cdot \frac{498}{193} - 3 \cdot \frac{1}{193} = \frac{1}{100} \\
    31 \cdot \frac{498}{193} + 2 \cdot \frac{1}{193} = \frac{64}{100}
\end{cases} \quad \rightarrow \text{Optimal}
\]

- Feasible solution: \(x^* = \begin{bmatrix} 0 \\ 2 \end{bmatrix}\) optimal?

- \(y_1^* = 0\) if \(\sum_{j=1}^{n} a_{ij} x_j^* < b_i\)

\[
\begin{cases} 
    50 \cdot 0 + 31 \cdot 2 = 62 < 250 \\
    -3 \cdot 0 + 2 \cdot 2 = 4
\end{cases} \quad \rightarrow y_1 = 0
\]

- \(\sum_{i=1}^{m} a_{ij} y_i^* = c_j\) if \(x_j^* > 0\)

\[
\begin{cases} 
    31y_1 + 2y_2 = \frac{64}{100} \\
    y_1 = 0
\end{cases} \quad y^* = \begin{bmatrix} 0 \\ \frac{32}{100} \end{bmatrix}
\]

- \(y_i^* \geq 0\) forall \(i = 1, \ldots, m \rightarrow \text{OK}\)
\[ \sum_{i=1}^{m} a_{ij} y_i^* \geq c_j \text{ for all } j = 1, \ldots, n \]

\[ \begin{align*}
50 \cdot 0 - 3 \cdot \frac{32}{100} &= -\frac{96}{100} < 1 \\
31 \cdot 0 + 2 \cdot \frac{32}{100} &= \frac{64}{100}
\end{align*} \rightarrow \text{Not Optimal} \]
5 Asset Pricing and Arbitrage

5.1 Pricing and Hedging of Options

We first start with a description of some of the well-known financial options.

**European call option:** is a contract with the following conditions

- At a prescribed time in the future, known as the *expiration date*, the holder of the option has the right, but not the obligation to purchase a prescribed asset, known as the *underlying*, for a prescribed amount, known as the *strike price* or *exercise price*.

**European put option:** is similar, except that it confers the right to sell the underlying asset.

**American option:** is like a European option, but it can be exercised anytime before the expiration date.

How can we determine the price of the options?

We start with an example and then we formalize the problem.

- To find the *fair value of an option*, we need to solve a *pricing problem*.
- Option pricing problems are often solved using the following strategy:
  - We try to determine a portfolio of assets with known prices which, if updated properly through time, will produce the same payoff (value) as the option.
  - Here a portfolio is the portfolio resulting from a decision of buying or selling stocks and/or selling or borrowing money
  - Since the portfolio and the option will have the same eventual payoffs, we conclude that they must have the same value today (otherwise, there is arbitrage) and we can therefore obtain the price of the option.
  - A portfolio of other assets that produces the same payoff as a given financial instrument is called a *replicating portfolio* (or a *hedge*).
  - Finding the right portfolio, of course, is not always easy and leads to a *replication (or hedging) problem*.

Let us consider a simple example to illustrate these ideas:

- We know that one share of stock $\alpha$ is currently valued at €40
- Assume that the value of $\alpha$ will either double or halve with equal probabilities (random values):

$$S_0=40 \quad S_1(u)=80 \quad S_1(d)=20$$

- Today, we purchase a European call option to buy one share of stock $\alpha$ for €50 a month from today. What is the fair price of this option?

Assumptions
• we can borrow or lend money with no interest between today and next month
• we can buy or sell any amount of the $\alpha$ stock without any commissions
• assume that $\alpha$ will not pay any dividends within the next month.

To solve the option pricing problem, we consider the following hedging problem:

• Can we form a portfolio of the underlying stock (bought or sold) and cash (borrowed or lent) today, such that the payoff of the portfolio at the expiration date of the option will match the payoff of the option?
• Note that the option payoff will be $\mathbb{E}30$ if the price of the stock goes up and $\mathbb{E}0$ if it goes down.
• Assume this portfolio has $\Delta$ shares of $\alpha$ and $B$ cash.

This portfolio would be worth $40\Delta + B$ today. Next month, payoffs for this portfolio will be:

\[ P_0 = 40\Delta + B \]
\[ \begin{align*}
80\Delta + B &= P_1(u) \\
20\Delta + B &= P_1(d)
\end{align*} \]

Let us choose $\Delta$ and $B$ such that:

\[ 80\Delta + B = 30 \]
\[ 20\Delta + B = 0 \]

so that the portfolio replicates the payoff of the option at the expiration date.

• This gives $\Delta = \frac{1}{2}$ and $B = -10$, which is the hedge we were looking for.
• This portfolio is worth $P_0 = 40\Delta + B = \mathbb{E}10$ today, therefore, the fair price of the option must also be $\mathbb{E}10$.

### 5.2 Replication Problem

In this Section we generalize the example of the previous paragraph.

• Let $S_0$ be the current price of the underlying security
• Assume that there are two possible outcomes at the end of the period:

\[ S_1^u = S_0 \cdot u \]
\[ S_1^d = S_0 \cdot d \]

(Assume $u > d$.)

• We also assume that there is a fixed interest paid on cash borrowed or lent at rate $r$ for the given period.

\[ R = 1 + r \]

• We consider a derivative security which has payoffs of $C_1^u$ and $C_1^d$ in the up and down states respectively.
• Considering the Strike Price $SP$, the payoff of an European Call Option can be computed as:

$C^u_1 = \max\{0, S^u_1 - SP\}$ \quad $C^d_1 = \max\{0, S^d_1 - SP\}$

• The replication problem considers a portfolio of $\Delta$ shares of the underlying and $B$ cash.

• For what values of $\Delta$ and $B$ does this portfolio have the same payoffs at the expiration date as the derivative security?

• We need to solve the following system of 2 equations and 2 variables ($\Delta$ and $B$):

\[
\begin{align*}
\Delta \cdot S_0 \cdot u + B \cdot R &= C^u_1 \\
\Delta \cdot S_0 \cdot d + B \cdot R &= C^d_1
\end{align*}
\]

• This portfolio is worth $S_0 \cdot \Delta + B$ today, which corresponds to the fair price of the option as well:

$C_0 = \frac{C^u_1 - C^d_1}{u - d} + \frac{u \cdot C^d_1 - d \cdot C^u_1}{R(u - d)}$  

$= \frac{1}{R} \left[ \frac{R - d}{u - d} C^u_1 + \frac{u - R}{u - d} C^d_1 \right]$  

• In the same way we can compute the hedge

$\Delta = \frac{C^u_1 \cdot R - C^d_1 \cdot R}{S_0 \cdot u \cdot R - S_0 \cdot d \cdot R}$  

$B = \frac{S_0 \cdot u \cdot C^d_1 - S_0 \cdot d \cdot C^u_1}{S_0 \cdot u \cdot R - S_0 \cdot d \cdot R}$

**Risk-Neutral Probabilities**

Let us define:

$pu = \frac{R - d}{u - d}$

and

$pd = \frac{u - R}{u - d}$

• Note that we must have $d < R < u$ to avoid arbitrage opportunities

• An immediate consequence of this observation is that both $pu > 0$ and $pd > 0$

• Noting also that $pu + pd = 1$ one can interpret $pu$ and $pd$ as probabilities

• These are the so-called risk-neutral probabilities (RNPs) of up and down states, respectively

• Note that they are completely independent from the actual probabilities of these states.
• The price of any derivative security can be calculated as the present value of the expected value of its future payoffs where the expected value is computed using the risk-neutral probabilities.

\[ C_0 = \frac{1}{R} (p_u \cdot C^u_1 + p_d \cdot C^d_1) \]

\[ P_0 = \frac{1}{R} (p_u \cdot P^u_1 + p_d \cdot P^d_1) \]

and we also have:

\[ S_0 = \frac{1}{R} (p_u \cdot S^u_1 + p_d \cdot S^d_1) \]

**Example**

In our example above \( u = 2, \ d = \frac{1}{2} \) and \( r = 0 \) so that \( R = 1 \). Therefore:

\[ p_u = \frac{1}{3} \quad \text{and} \quad p_d = \frac{2}{3} \]

As a result, we have

\[ S_0 = 40 = \frac{1}{R} (p_u \cdot S^u_1 + p_d \cdot S^d_1) = \frac{1}{3} 80 + \frac{2}{3} 20 \]

\[ C_0 = 10 = \frac{1}{R} (p_u \cdot C^u_1 + p_d \cdot C^d_1) = \frac{1}{3} 30 + \frac{2}{3} 0 \]

• Using risk neutral probabilities we can also price other derivative securities

• For example, consider a European put option a stock struck at 60€ (this is another way to say “with a strike price of 60€”) and with the same expiration date as the call of the example.

\[ P^u_1 = \max \{0, 60 - 80\} = 0 \]

\[ P^d_1 = \max \{0, 60 - 20\} = 40 \]

• Considering the Strike Price \( SP \), the payoff of an European Put Option can be computed as:

\[ P^u_1 = \max \{0, SP - S^u_1\} \quad P^d_1 = \max \{0, SP - S^d_1\} \]

• We can easily compute:

\[ P_0 = \frac{1}{R} (p_u \cdot P^u_1 + p_d \cdot P^d_1) = \frac{1}{3} 0 + \frac{2}{3} 40 = \frac{80}{3} \]

**Generalization from a binomial setting to a more general setting of \( m \) possible states and \( n \) different securities**

Let:

\[ \Omega = \{\omega_1, \omega_2, \ldots, \omega_m\} \]

be the (finite) set of possible future states.

For securities \( S^i \) with \( i = 1, \ldots, n \):
• let $S_i^1(\omega_j)$ denote the price of this security $i$ at time 1

For example, the following set of 3 ($m = 3$) of different prices of 2 ($n = 3$) different securities at a future date:

<table>
<thead>
<tr>
<th>State $\omega_1$ ($j = 1$)</th>
<th>Security 1 ($i = 1$)</th>
<th>Security 2 ($i = 2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_1^1(\omega_1)$ = 10</td>
<td>$S_2^1(\omega_1)$ = 30</td>
<td></td>
</tr>
<tr>
<td>State $\omega_2$ ($j = 2$)</td>
<td>$S_1^1(\omega_2)$ = 15</td>
<td>$S_2^1(\omega_2)$ = 15</td>
</tr>
<tr>
<td>State $\omega_3$ ($j = 3$)</td>
<td>$S_1^1(\omega_3)$ = 13</td>
<td>$S_2^1(\omega_3)$ = 25</td>
</tr>
</tbody>
</table>

• let $S_0^i$ denote the current (time 0) price of security $S^i$

For example, the current prices are $S_0^1 = 6$ and $S_0^2 = 20$.

• We use $i = 0$ for the "riskless" security that pays the interest rate $r \geq 0$ between time 0 and time 1
• It is convenient to assume that $S_0^0 = 1$ and that $S_0^i(\omega_j) = R = 1 + r, \forall j$.

**Definition** A risk-neutral probability measure is a vector of positive numbers $(p_1, p_2, \ldots, p_m)$ such that:

$$\sum_{j=1}^{m} p_j = 1$$

and for every security $S^i$, $i = 0, \ldots, n$:

$$S_0^i = \frac{1}{R} \left( \sum_{j=1}^{m} p_j S_i^1(\omega_j) \right)$$

### 5.3 The Fundamental Theorem of Asset Pricing

In this section we state the first fundamental theorem of asset pricing and prove it for finite $\Omega$ using the theory of linear programming.

We recall that:

**Definition** An arbitrage is a trading strategy:

• that has a positive initial cash flow and has no risk of a loss later (type A)
• that requires no initial cash input, has no risk of a loss, and a positive probability of making profits in the future (type B).

**Theorem 5.1 (Fundamental Theorem of Asset Pricing).** A risk-neutral probability measure exists if and only if there is no arbitrage.

**Proof.**

• We assume that the state space $\Omega$ is finite.

• We assume without loss of generality that every state has a positive probability of occurring (since states that have no probability of occurring can be removed)

• Given the current prices $S_0^i$ and the future prices $S_1^i(\omega_j)$ in each state $\omega_j$, for securities 0 to $n$, consider the following linear program with variables $x_i$, for $i = 0, \ldots, n$: 

---

32
\[
\begin{align*}
\min \sum_{i=0}^{n} S_i^0 \cdot x_i \\
\sum_{i=0}^{n} S_i^j (\omega_j) \cdot x_i \geq 0 \quad j = 1, \ldots, m \\
x_0 \leq 0 \\
x_i \geq 0 \quad i = 1, \ldots, n
\end{align*}
\]

- Note that type-A arbitrage corresponds to a feasible solution to this LP with a negative objective value.
- Since \( x_i = 0 \ (i = 0, \ldots, n) \) is a feasible solution, the optimal objective value is always non-positive.
- Furthermore, since all the constraints are homogeneous (0 RHS), if there exists a feasible solution such that

\[\sum_{i=0}^{n} S_i^0 \cdot x_i < 0\]

(this corresponds to type-A arbitrage), the problem is unbounded.
- In other words, there is no type-A arbitrage if and only if the optimal objective value of this LP is 0.

In the previous example this models reads as follows (with \( r \) at 4%):

\[
\begin{align*}
\min \quad x_0 + 6x_1 + 20x_2 \\
1.04x_0 + 10x_1 + 30x_2 \geq 0 \\
1.04x_0 + 15x_1 + 15x_2 \geq 0 \\
1.04x_0 + 13x_1 + 25x_2 \geq 0 \\
x_0 \leq 0, x_1, x_2 \geq 0
\end{align*}
\]

and the solution \( x_0 = -27, x_1 = 1 \) and \( x_2 = 1 \) is feasible.

\[
\begin{align*}
1.04 \cdot (-27) + 10 \cdot (1) + 30 \cdot (1) &= 11.920 \geq 0 \\
1.04 \cdot (-27) + 15 \cdot (1) + 15 \cdot (1) &= 1.920 \geq 0 \\
1.04 \cdot (-27) + 13 \cdot (1) + 25 \cdot (1) &= 9.920 \geq 0
\end{align*}
\]

Its value is \(-27 + 6 + 20 = -1\). In this case we have the type-A arbitrage! and the model is unbounded.

If we consider instead the following prices: \( S_0^1 = 13 \) and \( S_0^2 = 23 \):

\[
\begin{align*}
\min \quad x_0 + 13x_1 + 23x_2 \\
1.04x_0 + 10x_1 + 30x_2 \geq 0 \\
1.04x_0 + 15x_1 + 15x_2 \geq 0 \\
1.04x_0 + 13x_1 + 25x_2 \geq 0 \\
x_0 \leq 0, x_1, x_2 \geq 0
\end{align*}
\]

The solution \( x_0 = 0, x_1 = 0 \) and \( x_2 = 0 \) is feasible. Its value is 0. Now we have to prove that this solution is optimal. By complementing \( x_0 = -\bar{x}_0 \) we get the standard form:
\[
\begin{align*}
\min & \quad -\bar{x}_0 + 13x_1 + 23x_2 \\
& \quad -1.04\bar{x}_0 + 10x_1 + 30x_2 \geq 0 \\
& \quad -1.04\bar{x}_0 + 15x_1 + 15x_2 \geq 0 \\
& \quad -1.04\bar{x}_0 + 13x_1 + 25x_2 \geq 0 \\
\bar{x}_0, x_1, x_2 & \geq 0
\end{align*}
\]

(72)

Since the constraints are homogeneous and \(x_0 = 0, x_1 = 0\) and \(x_2 = 0\) is feasible it is sufficient to find a feasible dual solution:

\[
\begin{align*}
\max & \quad 0 \cdot y_1 + 0 \cdot y_2 + 0 \cdot y_3 \\
& \quad 1.04y_1 + 1.04y_2 + 1.04y_3 \geq 1 \\
& \quad 10y_1 + 15x_1 + 13x_2 \leq 13 \\
& \quad 30y_1 + 15x_1 + 25x_2 \leq 23 \\
y_1, y_2, y_3 & \geq 0
\end{align*}
\]

(77) \quad (78) \quad (79) \quad (80) \quad (81)

and a feasible solution is

\[
p_1^* = 0.5, p_2^* = 0.5, p_3^* = 0.000
\]

\[
1.04 \cdot (0.5) + 1.04 \cdot (0.5) + 1.04 \cdot (0) = 1.040 \geq 1 \\
10 \cdot (0.5) + 15 \cdot (0.5) + 13 \cdot (1) = 12.500 \leq 13 \\
30 \cdot (0.5) + 15 \cdot (0.5) + 25 \cdot (1) = 22.500 \leq 23
\]

(82) \quad (83) \quad (84)

• Suppose that there is no type-A arbitrage.

• Then, there is no type-B arbitrage if and only if all constraints are tight for all optimal solutions of the previous LP since every state has a positive probability of occurring.

• Consider the dual of the LP:

\[
\begin{align*}
\max & \quad \sum_{j=1}^{m} 0 \cdot p_j \\
\sum_{j=1}^{m} (1 + r) \cdot p_j & \geq 1 \\
\sum_{j=1}^{m} S_i^1(\omega_j) \cdot p_j & \leq S_i^0 \quad i = 1, \ldots, n \\
p_j & \geq 0 \quad j = 1, \ldots, m
\end{align*}
\]

(85) \quad (86) \quad (87) \quad (88)

• Since the dual has a constant (0) objective function, any dual feasible solution is also dual optimal.

• When there is no type-A arbitrage the LP has an optimal solution, and the Strong Duality Theorem indicates that the dual must have a feasible solution

• If there is no type-B arbitrage also, the complementary slackness theorem indicates that there exists a feasible (and therefore optimal) dual solution \(p^*\) such that \(p > 0\)
• From the dual constraint corresponding to \( i = 0 \), we have that

\[
\sum_{i=1}^{n} p_i^* = \frac{1}{R}
\]

• Multiplying \( p^* \) by \( R \) one obtains a risk-neutral probability distribution.

• Therefore, the "no arbitrage" assumption implies the existence of RNPs

• The converse direction is proved in an identical manner.

For our example and prices 13 and 23, the dual problem reads as follows:

\[
\begin{align*}
\text{max} & \quad 0 \cdot y_1 + 0 \cdot y_2 + 0 \cdot y_3 \\
& \quad 1.04y_1 + 1.04y_2 + 1.04y_3 \geq 1 \\
& \quad 10y_1 + 15y_2 + 13y_3 \leq 13 \\
& \quad 30y_1 + 15y_2 + 25y_3 \leq 23 \\
& \quad y_1, y_2, y_3 \geq 0
\end{align*}
\]

and the optimal solution is

\[
p_1^* = 0.285, p_2^* = 0.677, p_3^* = 0.000
\]

. And the risk neutral probabilities are

\[
0.296; 0.704; 0.000
\]

obtained by multiplying \( p^* \) by \( R \).
6 Lagrangian Relaxation

Problem $P$ to relax: (purely binary problem, but results extend to ILP)

\[
Z(P) = \max \sum_{j=1}^{n} v_j x_j
\]  
(94)
\[
\sum_{j=1}^{n} a_{ij} x_j \leq b_i \quad i = 1, \ldots, m
\]  
(95)
\[
\sum_{j=1}^{n} d_{kj} x_j = e_k \quad k = 1, \ldots, l
\]  
(96)
\[
x_j \in \{0, 1\} \quad j = 1, \ldots, n
\]  
(97)

or equivalently

\[
Z(P) = \max \mathbf{v}^\top \mathbf{x}
\]  
(98)
\[
A \mathbf{x} \leq \mathbf{b}
\]  
(99)
\[
D \mathbf{x} = \mathbf{e}
\]  
(100)
\[
\mathbf{x} \in \{0, 1\}
\]  
(101)

(A) Inequalities

1. Select a set of inequalities constraints and add a non-negative term to the objective function:

\[
\max \sum_{j=1}^{n} v_j x_j + \sum_{i=1}^{m} u_i (b_i - \sum_{j=1}^{n} a_{ij} x_j)
\]  
(102)

with $u_i \geq 0$, $\forall i$. Lagrangian multiplier (valid upper bound).

2. Eliminate the selected constraints. Lagrangian Relaxation:

\[
Z(L(P, \mathbf{u})) = \sum_{i=1}^{m} u_i b_i + \max \sum_{j=1}^{n} \hat{v}_j x_j
\]  
(103)
\[
\sum_{j=1}^{n} d_{kj} x_j = e_k \quad k = 1, \ldots, l
\]  
(104)
\[
x_j \in \{0, 1\} \quad j = 1, \ldots, n
\]  
(105)

where $\hat{v}_j = v_j - \sum_{i=1}^{m} u_i a_{ij}$ $(j = 1, \ldots, n)$.

- If the optimal solution \( \mathbf{x}^* \) to $L(P)$ is feasible for $P$ then it is not necessarily optimal for $P$: it is optimal if the two objective functions have the same value, i.e., if

\[
\sum_{i=1}^{m} u_i (b_i - \sum_{j=1}^{n} a_{ij} x_j) = 0
\]

($\iff$ if, $\forall i$, either $u_i$ is 0 or constraint $i$ is tight)
(B) Equalities

1. Select a set of equality constraints and add a null term to the objective function:

$$\max \sum_{j=1}^{n} v_j x_j + \sum_{k=1}^{l} u_k (e_k - \sum_{j=1}^{n} d_{kj} x_j)$$ (106)

with $u_k \leq = \geq 0, \forall k$. **Lagrangian multiplier** (problem unchanged).

2. Eliminate the selected constraints. **Lagrangian Relaxation**:

$$Z(L(P, u)) = \sum_{k=1}^{l} u_k e_k + \max \sum_{j=1}^{n} \hat{v}_j x_j$$ (107)

$$\sum_{j=1}^{n} a_{ij} x_j \leq b_i \quad i = 1, \ldots, m$$ (108)

$$x_j \in \{0, 1\} \quad j = 1, \ldots, n$$ (109)

where $\hat{v}_j = v_j - \sum_{k=1}^{l} u_k d_{kj}$ ($j = 1, \ldots, n$).

- If the optimal solution $x^*$ to $L(P)$ is feasible for $P$ then it optimal for $P$.

**Graphical representation**

Ineq: unfeasible solution – feasible non optimal –feasible and optimal solution solution

![Graphical representation](image)

Eq: unfeasible solution –feasible and optimal solution solution

![Graphical representation](image)
(C) Lagrangian Dual Problem \( D_L \)

(Ineq.) Lagrangian dual problem: find \( \bar{u} \) such that \( Z(L(P, \bar{u})) = \min_{u \geq 0} Z(L(P, u)) \).

(Eq.) Lagrangian dual problem: find \( \bar{u} \) such that \( Z(L(P, \bar{u})) = \min_u Z(L(P, u)) \).

- In case \( Z(D_L) > Z(P) \) then we have Lagrangian Duality Gap between \( D_L \) and \( P \).

The following theorems hold (written for the inequalities case but valid also for the equalities case)

**Theorem 6.1 (Lagrangian Weak Duality).** The optimal objective function value \( Z(P) \) of problem \( P \) is always smaller or equal to the optimal objective function value \( Z(L(P, u)) \) of problem \( L(P, u) \):

\[
Z(P) \leq Z(L(P, u)), \quad \forall u \geq 0.
\]

**Proof.** Let \( x^* \) be the optimal solution of the problem \( P \), then \( x^* \) is also a feasible solution of problem \( L(P) \) for each \( u \geq 0 \):

\[
c^\top x^* \leq Z_L(x^*) \quad c^\top x^* + u(b - Ax^*) \leq Z_L(P, u)
\]

the quantity \((b - Ax^*)\) is \( \geq 0 \), then:

\[
c^\top x^* \leq Z(L(P, u))
\]

proving that \( Z(L(P, u)) \geq Z(P) \)

In case of inequalities it is clear that there is one point in which \( Z(L(P, u)) \geq Z(P) \) within the feasible region of \( P \). In case of equalities instead within the feasible region of \( P \), we have \( Z(L(P, u)) = Z(P) \), but since we have relaxed some constraints and enlarged the feasible region of \( P \) then \( Z(L(P, u)) \geq Z(P) \).

**Theorem 6.2 (Lagrangian Strong Duality).** Let \( \bar{x} \) be the optimal solution value of \( L(P, \bar{u}) \) for a specific \( \bar{u} \geq 0 \). If \( \bar{x} \) and \( \bar{u} \) are such that: \( \bar{x} \) is feasible for \( P \) (i.e., \( A\bar{x} \geq b \)) and \( u(A\bar{x} - b) = 0 \), then \( \bar{x} \) is the optimal solution of \( P \).

**Proof.** We start by demonstrating that \( \bar{x} \) is an optimal solution of \( P \). Since \( \bar{x} \) is a feasible solution for \( P \) we have:

\[
c^\top \bar{x} \leq Z(P).
\]

For the Lagrangian Weak Duality we have:

\[
Z(P) \leq Z(L(P, \bar{u})) = c^\top \bar{x} + u(b - A\bar{x})
\]

We can then write:

\[
Z(P) \leq c^\top \bar{x} \leq Z(P),
\]

or in other words \( Z(P) = Z(L(P, \bar{u})) \)

- We demonstrate now the \( Z(D_L) = Z(L(P, \bar{u})) \). For construction of the problem \( D_L \) we have \( Z(D_L) \leq Z(L(P, \bar{u})) \) and \( Z(P) \leq Z(D_L) \); we have then:

\[
Z(D_L) = Z(P).
\]
Subgradient Optimization  Good multipliers are produced by the subgradient method. We consider the case of Lagrangian relaxation of inequalities:

- objective function: \( \max \sum_{j=1}^{n} v_j x_j + \sum_{i=1}^{m} u_i L_i \) \( (u_i \geq 0, \forall i) \)

where subgradient \( L_i = b_i - \sum_{j=1}^{n} a_{ij} x_j \) is the slack in the \( i \)-th relaxed constraint:

- if \( L_i < 0 \) (violated constraint) the term \( u_i L_i \) penalizes the objective function;
- if \( L_i > 0 \) (loose constraint) the term \( u_i L_i \) rewards the objective function.

- Algorithm:
  start with any \( u \), and a prefixed initial step \( t > 0 \);
  while halting condition is not satisfied do
  begin
    solve \( L(P, u) \);
    for \( i := 1 \) to \( m \) do \( u_i := \max(0, u_i - t L_i) \);
    correct \( t \);
  end

- Intuitively:
  - if \( L_i < 0 \) then \( u_i \) is too small and must increase;
  - if \( L_i > 0 \) then \( u_i \) is too large and must decrease;
  - if \( L_i = 0 \) then \( u_i \) is fine (optimality condition).

- Standard formula for the step updating:
  \[
  t = \vartheta \frac{Z(L(P, u)) - Z}{\sum_{i=1}^{m} L_i^2}
  \]

  where
  - \( Z \) = incumbent solution value;
  - \( \vartheta \) decreases with the number of iterations, e.g., start with \( \vartheta = 2 \); halve \( \vartheta \) after some iterations with no improvement.

- Two halting conditions:
  - \( \forall i \), either \( L_i = 0 \) or \( L_i > 0 \) and \( u_i = 0 \) (optimal multipliers);
  - maximum number of iterations reached.

- Non-monotone method: the best \( u \) value is stored.

Different possible kinds of relaxation

- \( F(P) \) = feasible region of \( P \)
- \( F_r(P) \) = feasible region of the relaxation \( r \) of \( P \)
7 Construction of an Index Fund

Different possibilities of portfolio management:

- **active portfolio management** tries to achieve superior performance by using forecasting techniques
- **passive portfolio management** avoids any forecasting techniques and rather relies on diversification to achieve a desired performance

There are 2 types of passive management strategies:

- **buy and hold**: assets are selected on the basis of some fundamental criteria and there is no active selling or buying of these stocks afterwards
- **indexing**: the goal is to choose a portfolio that mirrors the movements of a broad market population or a market index. Such a portfolio is called an index fund.

**Index fund – Definition**  Given a target population of $n$ stocks, one selects $q$ stocks and their weights in the index fund, to represent the target population as closely as possible.

In the last twenty years, an increasing number of investors, both large and small, have established index funds. The rising popularity of index funds can be justified both theoretically and empirically:

- **Market Efficiency**: If the market is efficient, no superior risk-adjusted returns can be achieved by stock picking strategies since the prices reflect all the information available in the marketplace. An index fund captures the efficiency of the market via diversification.
- **Empirical Performance**: Considerable empirical literature provides strong evidence that, on average, money managers have consistently underperformed the major indexes (luck is an explanation for good performance).
- **Transaction Cost**: Actively managed funds incur transaction costs, which reduce the overall performance of these funds. In addition, active management implies significant research costs.

Additional features:

- Here we take the point of view of a fund manager who wants to construct an index fund.
- Strategies for forming index funds involve choosing a broad market index as a proxy for an entire market, e.g. the Standard and Poor list of 500 stocks (S&P 500)
- A pure indexing approach consists in purchasing all the issues in the index, with the same exact weights as in the index.
- In most instances, this approach is impractical (many small positions) and expensive (rebalancing costs may be incurred frequently)

**A Large-Scale Deterministic Model**

- We propose a large-scale deterministic model for aggregating a broad market index of stocks into a smaller more manageable index fund.
- This approach will produce a portfolio that closely replicates the underlying market population.

We present a model that:

- clusters the assets into groups of similar assets
- selects one representative asset from each group to be included in the index fund portfolio
The model is based on the following data:

\[ \rho_{ij} = \text{similarity between stock } i \text{ and stock } j \]  

(110)

Typical features:
- \( \rho_{ii} = 1 \)
- \( \rho_{ij} \leq 1 \) for \( i \neq j \)
- \( \rho_{ij} \) is larger for more similar stocks
- An example of this is the correlation between the returns of stocks \( i \) and \( j \)

Decision variables:

\[
y_j = \begin{cases} 
1 & \text{if stock } j \text{ is in the index fund} \\
0 & \text{otherwise} 
\end{cases}
\]

\[
x_{ij} = \begin{cases} 
1 & \text{if } j \text{ is the most similar stock to stock } i \text{ in the index fund} \\
0 & \text{otherwise} 
\end{cases}
\]

\[
\max Z(IF) = \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_{ij} x_{ij} 
\]

(111)

\[
\sum_{j=1}^{n} y_j = q
\]

(112)

\[
\sum_{j=1}^{n} x_{ij} = 1 \quad i = 1, \ldots, n
\]

(113)

\[ x_{ij} \leq y_j \quad i = 1, \ldots, n, j = 1, \ldots, n 
\]

(114)

\[ x_{ij} \in \{0, 1\} \quad i = 1, \ldots, n, j = 1, \ldots, n 
\]

(115)

\[ y_j \in \{0, 1\} \quad j = 1, \ldots, n 
\]

(116)

- Interpret each of the constraints
- Explain why the objective of the model can be interpreted as selecting \( q \) stocks out of the population of \( n \) stocks so that the total loss of information is minimized

Once the model has been solved and a set of \( q \) stocks has been selected for the index fund, a weight \( w_j \) is calculated for each \( j \) in the fund based on a measure of the importance (\( V_i \)) of stock \( i \) in the market:

\[
w_j = \sum_{i=1}^{n} V_i x_{ij}
\]

(117)

One possibility is to use for \( V_i \) the market value of stock \( i \) (Market value is commonly used to refer to the market capitalization of a publicly-traded company, and is obtained by multiplying the number of its shares by the current share price.).

- So \( w_j \) is the total market value of the stocks represented by stock \( j \) in the fund
- The fraction of the index fund to be invested in stock \( j \) is proportional to the stocks weight \( w_j \)
- \[
\frac{w_j}{\sum_{f=1}^{n} w_f}
\]
- Note that, instead of the objective function used, one could have used an objective function that takes the weights \( w_j \) directly into account \[ \sum_{i=1}^{n} \sum_{j=1}^{n} V_i \rho_{ij} x_{ij} \]
Solution strategies

- Branch-and-bound is a natural candidate for solving the model, however the formulation can be very large. Indeed, for the S&P 500, there are 250,000 variables $x_{ij}$ and 250,000 constraints $x_{ij} \leq y_j$

- The linear programming relaxation needed to get upper bounds in the branch-and-bound algorithm is difficult to solve

- Instead of the linear programming relaxation, one can use the Lagrangian relaxation

$$\max \ Z(L(IF, u)) = \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_{ij} x_{ij} + \sum_{i=1}^{n} u_i (1 - \sum_{j=1}^{n} x_{ij})$$

(118)

$$\sum_{j=1}^{n} y_j = q$$

(119)

$$x_{ij} \leq y_j \quad i = 1, \ldots, n, j = 1, \ldots, n$$

(120)

$$x_{ij} \in \{0, 1\} \quad i = 1, \ldots, n, j = 1, \ldots, n$$

(121)

$$y_j \in \{0, 1\} \quad j = 1, \ldots, n$$

(122)

where $u = (u_1, \ldots, u_n)$ is the vector of the Lagrangian Multipliers for the constraints (113)

The objective function of $L(IF, u)$ may be equivalently stated as:

$$\max \ Z(L(IF, u)) = \sum_{j=1}^{n} C_j y_j + \sum_{i=1}^{n} u_i$$

(123)

Let:

$$(\rho_{ij} - u_i)^+ = \begin{cases} (\rho_{ij} - u_i) & \text{if } \rho_{ij} - u_i \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

and

$$C_j = \sum_{i=1}^{n} (\rho_{ij} - u_i)^+$$

(124)

Remark. Model ((118)-(122)) can be equivalently rewritten as follow:

$$\max \ Z(L(IF, u)) = \sum_{j=1}^{n} C_j y_j + \sum_{i=1}^{n} u_i$$

(125)

$$\sum_{j=1}^{n} y_j = q$$

(126)

$$y_j \in \{0, 1\} \quad j = 1, \ldots, n$$

(127)

- Where the $x$ variables has been projected away, since the optimal solution depends only on the $y$ variables:

$$x_{ij} = y_j \quad \text{or} \quad x_{ij} = 0$$

Remark. In an optimal solution of the Lagrangian relaxation, $y_j$ is equal to 1 for the $q$ largest values of $C_j$, and the remaining $y_j$ are equal to 0.
• Once the optimal solution in term of the \( y \) variables has been computed then the \( x \) variables can be computed: \( x_{ij} = 1 \) if \( \rho_{ij} - u_i \geq 0 \) and \( y_j = 1 \) (0 otherwise).

• it worth mentioning that in the solution of the Lagrangian relaxation more one stock \( i \) can be assigned to each stock \( j \) in the index fund (or zero).

• Interestingly, the set of \( q \) stocks corresponding to the \( q \) largest values of \( C_j \) can also be used to build a heuristic solution for the model IF. Specifically, construct an index fund containing these \( q \) stocks and assign each stock \( i = 1, \ldots, n \) to the most similar stock in this fund. This solution is feasible to the model, although not necessarily optimal.

• So for any vector \( u \), we can compute quickly both a lower bound and an upper bound on the optimum value of the model IF.

How does one minimize \( L(IF, u) \)?

• Since \( L(u) \) is non-differentiable and convex, one can use the subgradient method

• At each iteration, a revised set of Lagrange multipliers \( u \) and an accompanying lower bound and upper bound to the model are computed

• At each iteration constraints (113) can be violated (some stock can be assigned to more than one stock or to zero). Accordingly a sub-gradient optimisation method can be used (See the appendix for further details).

• The algorithm terminates when these two bounds match or when a maximum number of iterations is reached

• It can be proven that \( \min Z(L(IF, u)) \) is equal to the optimal value of model \( IF \), i.e., \( Z(IF) \).

• To solve the integer program to optimality, one can also use a branch-and-bound algorithm, using as upper bound \( Z(L(IF, u)) \) for pruning the nodes.