2 MVO and factor models

2.1 Review of MVO

The typical notation for MVO is:

- n risky assets,
- r_i random rate of return for asset i (observed return: $r_{it} = \frac{P_t - P_{t-1}}{P_{t-1}}$)
- $\mu_i = \mathbb{E}[r_i], \quad \sigma_i^2 = \text{var}(r_i), \quad \sigma_{ij} = \text{cov}(r_i, r_j)$ for $i \neq j$,
- R is the expected return goal,
- x_i is the proportion of wealth invested in asset i (sometimes w is used for the weights),
- $r_p = r_1 x_1 + r_2 x_2 + \ \ldots \ + r_n x_n =$ the portfolio return (r.v.),
- $\mathbb{E}[r_p] = \mu_1 x_1 + \mu_2 x_2 + \ \ldots \ + \mu_n x_n =$ expected portfolio return,
- $\sigma_p^2 = \text{var}(r_p) = \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} x_i x_j \qquad$ (note that $\sigma_{ii} = \sigma_i^2$).

Some of these parameters can be estimated from raw data as follows:

- N is the number of observations,
- Geometric mean: $\mu_i = \left(\prod_{t=1}^{N} (1 + r_{it}) \right)^{1/N} - 1$,
- Arithmetic mean: $\bar{r}_i = \frac{1}{N} \sum_{t=1}^{N} r_{it}$,
- Variance: $\sigma_i^2 = \frac{1}{N} \sum_{t=1}^{N} (r_{it} - \mathbb{E}[r_i])^2$,

  If we only have a sample of the population, we use the sample variance:

  $\rightarrow$ Sample variance: $\sigma_i^2 = \frac{1}{N-1} \sum_{t=1}^{N} (r_{it} - \bar{r}_i)^2$,
- Covariance: $\sigma_{ij} = \frac{1}{N} \sum_{t=1}^{N} (r_{it} - \mathbb{E}[r_i])(r_{jt} - \mathbb{E}[r_j])$,

  If we only have a sample of the population, we use the sample covariance:

  $\rightarrow$ Sample covariance: $\sigma_{ij} = \frac{1}{N-1} \sum_{t=1}^{N} (r_{it} - \bar{r}_i)(r_{jt} - \bar{r}_j)$,
- Correlation between 2 assets: $\rho_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j}$

MVO Version 1 – Minimize risk

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_{ij} x_i x_j \\ \text{s.t.} \quad & \sum_{i=1}^{n} \mu_i x_i \geq R \\ & \sum_{i=1}^{n} x_i = 1 \\ & (x_i \geq 0, \quad i = 1,\ ...,\ n) \end{aligned} \qquad \iff \qquad \begin{aligned} \min_{\boldsymbol{x}} \quad & x^T Q\ x \\ \text{s.t.} \quad & \mu^T x \geq R \\ & 1^T x = 1 \\ & (x_i \geq 0, \quad i = 1,\ ...,\ n) \end{aligned}$$

Where we can use matrix notation to express the problem:

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix}, \quad \boldsymbol{Q} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & ... & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & ... & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & ... & \sigma_n^2 \end{bmatrix}, \quad \mathbf{1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix},$$

2.2 MVO Limitations

- Single-period static model (not forward looking)
- Deterministic (we must capture uncertainty implicitly).
- Only the first 2 moments matter (1st moment = mean, 2nd moment = variance)

  $\rightarrow$ this naively assumes the distribution is symmetric.
- It considers deviations from the expected returns as risk, even when the price movement might be to our advantage.
- Sensitive to parameter estimation (we will explore ways to mitigate this throughout the course).
- MVO can produce portfolios that are over-concentrated, i.e., only a few assets i have $x_i \neq 0$.
- MVO does not account for transaction costs.

2.3 Other MVO versions

MVO Version 2 – Maximize return

$$\begin{aligned}
\max_{\boldsymbol{x}} \quad & \sum_{i=1}^{n} \mu_i x_i \\
\text{s.t.} \quad & \sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_{ij} x_i x_j \le \epsilon^2 \\
& \sum_{i=1}^{n} x_i = 1 \\
& (x_i \ge 0, \quad i = 1,\ \ldots,\ n)
\end{aligned}$$

MVO Version 3 – Mixed

$$\begin{aligned}
\max_{\boldsymbol{x}} \quad & \sum_{i=1}^{n} \mu_i x_i - \lambda \sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_{ij} x_i x_j \\
\text{s.t.} \quad & \sum_{i=1}^{n} x_i = 1 \\
& (x_i \ge 0, \quad i = 1,\ \ldots,\ n)
\end{aligned}$$

Typical MVO constraints

The budget constraint: $\sum_{i=1}^{n} x_i = 1^T x = 1$ ensures that 100% of our budget is spent on assets. We wish to use all our available resources and we cannot exceed this amount.

No short-selling allowed: Adding the constraint $x_i \ge 0,$ for $i = 1,\ \ldots,\ n$ ensures that we can only take a "long" (positive) position on each asset. In other words, we are buying the asset.

Note: If this constraint is not present it means that we can short-sell the asset. We are borrowing the asset and selling it immediately, using the additional cash to invest it in other assets. We must returned the borrowed asset at some point in the future.

How many parameter estimations are required?

$\mu_i \sim n$

$\sigma_i^2 \sim n$

$\sigma_{ij} \quad (i \ne j) \sim \frac{n^2-n}{2} = \frac{n(n-1)}{2}$

So the total number of estimations is $2n + \frac{n(n-1)}{2} = O(n^2)$

e.g., $n = 500,\ > 120,000$ estimations, this is a lot!

How can we reduce the number of estimations required?

2.4 Factor Models

Factor models are very popular, both in academia and in the industry. Factors are the drivers that explain the asset returns. Typical factors are derived from:

- Market indices, such as the S&P 500

- Portfolios constructed by aggregating several assets (e.g., Fama–French model)
- Principal component analysis (PCA)
- Other economic indicators

2.4.1 Generic single-factor model

$$r_i = \alpha_i + \beta_i f_m + \varepsilon_i$$

- r_i: Random return of asset i
- α_i and β_i: Regression coefficients
- f_m: the factor (e.g., the market return)
- $\varepsilon_i \sim \mathcal{N}(0, \sigma^2_{\varepsilon_i})$ is the random noise (pertains to the idiosyncratic risk)

Single-Factor Model: Ideal environment:

1. $\text{cov}(f_m, \varepsilon_i) = 0$
2. $\text{cov}(\varepsilon_i, \varepsilon_j) = 0 \quad \text{for } i \neq j$

The idea is to use the factor model to generate the parameters for MVO:

$$r_i = \alpha_i + \beta_i f_m + \varepsilon_i$$

1.
$$\begin{aligned}
\mu_i = \mathbb{E}[r_i] &= \mathbb{E}[\alpha_i + \beta_i f_m + \varepsilon_i] \\
&= \alpha_i + \beta_i \bar{f}_m, \\
&\rightarrow \ \bar{f}_m = \mathbb{E}[f_m]
\end{aligned}$$

2.
$$\begin{aligned}
\sigma_i^2 = \text{var}(r_i) &= \mathbb{E}[(r_i - \mu_i)^2] \\
&= \mathbb{E}[(\alpha_i + \beta_i f_m + \varepsilon_i - \alpha_i - \beta_i \bar{f}_m)^2] \\
&= \mathbb{E}[(\beta_i (f_m - \bar{f}_m) + \varepsilon_i)^2] \\
&= \mathbb{E}[\beta_i^2 \ (f_m - \bar{f}_m)^2 + 2\beta_i \ (f_m - \bar{f}_m)\varepsilon_i + \varepsilon_i^2] \\
&= \beta_i^2 \ \text{var}(f_m) + 2\beta_i \text{cov}(f_m, \varepsilon_i) + \text{var}(\varepsilon_i) \\
&= \beta_i^2 \sigma_m^2 + \sigma_{\varepsilon_i}^2, \\
&\rightarrow \ \sigma_m^2 \text{ is the variance of the market,} \\
&\rightarrow \ \sigma_{\varepsilon_i}^2 \text{ is the idiosyncratic risk of the asset.}
\end{aligned}$$

3.
$$\begin{aligned}
\sigma_{ij} = \text{cov}(r_i, r_j) &= \mathbb{E}[(r_i - \mu_i)(r_j - \mu_j)] \\
&= \mathbb{E}[((\alpha_i + \beta_i f_m + \varepsilon_i) - (\alpha_i + \beta_i \bar{f}_m))((\alpha_j + \beta_j f_m + \varepsilon_j) - (\alpha_j + \beta_j \bar{f}_m))] \\
&= \mathbb{E}[(\beta_i (f_m - \bar{f}_m) + \varepsilon_i)(\beta_j (f_m - \bar{f}_m) + \varepsilon_j)] \\
&= \beta_i \beta_j \text{var}(f_m) + \beta_j \text{cov}(f_m, \varepsilon_i) + \beta_i \text{cov}(f_m, \varepsilon_j) + \text{cov}(\varepsilon_i, \varepsilon_j) \\
&= \beta_i \beta_j \sigma_m^2
\end{aligned}$$

The total number of estimations now is:

$$\begin{aligned}
&\alpha_i \sim n \quad \sigma_m^2 \sim 1 \quad \sigma_{\varepsilon_i}^2 \sim n \\
&\beta_i \sim n \quad \bar{f}_m \sim 1 \\
&\text{Total} = 3n + 2 = O(n)
\end{aligned}$$

Reducing the number of estimated parameters not only reduces computational cost, but also improves the stability of MVO (reduces overall uncertainty).

Example of a single-factor model: CAPM

$$r_i - r_f = \alpha_i + \beta_{im}(f_m - r_f) + \varepsilon_i$$

- r_i is the asset return (r.v.)
- r_f is the risk-free rate of return (e.g. U.S. Treasury Bills or bonds)
- $(f_m - r_f)$ is the market excess rate of return
- $\beta_i = \sigma_{im}/\sigma_m^2$

MVO under a single-factor model

$$\begin{aligned}
\min_x \quad & \sum_{i=1}^n \sigma_i^2 x_i^2 + \sum_{i \neq j} \sum \sigma_{ij} x_i x_j \\
\text{s.t.} \quad & \sum_{i=1}^n \mu_i x_i \geq R \\
& \sum_{i=1}^n x_i = 1 \\
& (x_i \geq 0, \quad i = 1,\ \dots,\ n)
\end{aligned}
\quad \Longleftrightarrow \quad
\begin{aligned}
\min_x \quad & \sum_{i=1}^n (\beta_i^2 \sigma_m^2 + \sigma_{\varepsilon_i}^2) x_i^2 + \sum_{i \neq j} \sum \beta_i \beta_j \sigma_m^2 x_i x_j \\
\text{s.t.} \quad & \sum_{i=1}^n \Big(\alpha_i + \beta_i \bar{f}_m\Big) x_i \geq R \\
& \sum_{i=1}^n x_i = 1 \\
& (x_i \geq 0, \quad i = 1,\ \dots,\ n)
\end{aligned}$$

Or, in vector notation

$$\begin{aligned}
\min_{\boldsymbol{x}} \quad & x^T Q\, x \\
\text{s.t.} \quad & \mu^T x \geq R \\
& 1^T x = 1 \\
& (x_i \geq 0, \quad i = 1,\ \dots,\ n)
\end{aligned}$$

where we must estimate Q and μ from the regression model.

Example:

We have the following historical prices of asset A. Calculate the returns using

$$r_{At} = \frac{P_t - P_{t-1}}{P_{t-1}}$$

Year	Price of A	Return of A
2010	80.00	-
2011	88.00	10%
2012	83.60	-5%
2013	89.87	7.5%

Now, consider a single factor to explain the returns of A.

Year	Returns of A	Returns of f
2011	10%	5%
2012	-5%	-3%
2013	7.5%	4.5%

We can perform an ordinary least squares (OLS) regression to calculate the regression parameters. OLS has a closed-form solution, which we can exploit by setting up a matrix $X = [1\ f]$.

$$X = \begin{bmatrix} 1 & 0.05 \\ 1 & -0.03 \\ 1 & 0.045 \end{bmatrix}$$

The first column of X is a column of ones to account for the intercept of regression. We can calculate the coefficients like this

$$\begin{bmatrix} \alpha_A \\ \beta_A \end{bmatrix} = (X^T X)^{-1} X^T r_A = \begin{bmatrix} 0.003 \\ 1.7842 \end{bmatrix}$$

Now that we have our coefficients, we can calculate the vector of residuals

$$\varepsilon = r_A - X \begin{bmatrix} \alpha_A \\ \beta_A \end{bmatrix}$$

We can calculate the <u>unbiased</u> variance of ε to get σ_ε^2 as follows

$$\sigma_\varepsilon^2 = \frac{1}{N-p-1} \sum_{t=1}^{N} (\varepsilon_t)^2.$$

where we divide by the appropriate degrees of freedom (DOF), $N-p-1$, where p is the number of factors in the model. Next, we calculate the sample variance of the factor

$$\sigma_f^2 = \frac{1}{N-1} \sum_{t=1}^{N} (f_t - \tilde{f})^2,$$

where we use $\tilde{f}$ to denote the arithmetic mean. The expected return of f, which we denote as $\bar{f}$, must be calculated using the geometric mean

$$\bar{f} = \Big(\prod_{t=1}^{N}(1 + f_{it})\Big)^{1/N} - 1.$$

Finally, we can use the factor model to compute the expected return and variance of our asset

$$\mu_A = \alpha_A + \beta_A \bar{f},$$
$$\sigma_A^2 = \beta_A^2 \sigma_f^2 + \sigma_{\varepsilon_A}^2.$$

2.4.2 Multi-factor models

Suppose there are p factors that drive the asset returns. We now have that

$$r_i = \alpha_i + \sum_{k=1}^{p} \beta_{ik} f_k + \varepsilon_i$$

where β_{ik} is the sensitivity of asset i to factor k

Multi-Factor Model: Ideal environment:

- $\text{cov}(f_i, \varepsilon_j) = 0$ for all $i,\ j$
- $\text{cov}(\varepsilon_i, \varepsilon_j) = 0$ for all $i \neq j$
- $\text{cov}(f_i, f_j) = 0$ for all $i \neq j$ $\quad \Rightarrow$ This condition is sometimes hard to satisfy

Estimate regression coefficients

We can estimate the regression coefficients of multiple assets with multiple factors with a single equation:

$$\begin{bmatrix} \alpha^T \\ \beta^T \end{bmatrix} = (X^T X)^{-1} X^T Y$$

where

- $Y \in \mathbb{R}^{N \times n}$ is the matrix of asset returns (each column is the timeseries of a single asset)
- $X = \begin{bmatrix} 1 & f_1 & \cdots & f_p \end{bmatrix} \in \mathbb{R}^{N \times (p+1)}$ is the matrix of predictors.
- $\alpha \in \mathbb{R}^n$ is the vector of intercepts for all n assets
- $\beta \in \mathbb{R}^{n \times p}$ is a matrix where each element, β_{ik}, is the factor loading of asset i corresponding to factor k.

Estimate MVO input parameters:

Scalar notation:

$$r_i = \alpha_i + \sum_{k=1}^{p} \beta_{ik} f_k + \epsilon_i$$

$$\mu_i = \alpha_i + \sum_{k=1}^{p} \beta_{ik} \bar{f}_k$$

$$\sigma_i^2 = \sum_{k=1}^{p} \beta_{ik}^2 \sigma_{f_k}^2 + \sigma_{\varepsilon_i}^2$$

$$\sigma_{ij} = \sum_{k=1}^{p} \beta_{ik} \beta_{jk} \sigma_{f_k}^2$$

Vector notation:

$$r = \alpha + \beta f + \epsilon$$

$$\mu = \alpha + \beta \bar{f}$$

$$Q = \beta F \ \beta^T + D$$

where

- $\beta \in \mathbb{R}^{n \times p}$ is the matrix of factor loadings (β_{ik}'s).
- $f \in \mathbb{R}^p$ is the vector of random factor returns.
- $\bar{f} \in \mathbb{R}^p$ is the vector of expected factor returns.
- $F \in \mathbb{R}^{p \times p}$ is the factor covariance matrix.
- $D \in \mathbb{R}^{n \times n}$ is the diagonal matrix of residual variances $\sigma_{\epsilon_i}^2$.

Example of a 3-factor model: The Fama–French 3-factor model

$$r_i - r_f = \alpha_i + \beta_{im}(f_m - r_f) + \beta_{is} SMB + \beta_{iv} HML + \varepsilon_i$$

- r_i is the asset return (r.v.)
- r_f is the risk-free rate of return (e.g. U.S. Treasury Bills or bonds)
- α_i is the active return (return over the benchmark), i.e., the y-intercept of excess return
- β_{im} is the market factor loading of asset$i$, i.e., the asset's exposure to the market risk
- $(f_m - r_f)$ is the excess market return
- β_{is} is the size factor loading of asset i, i.e., the asset's exposure to the size risk
- SMB is the *Small Minus Big* factor, computed as the average return of a portfolio composed of small cap stocks minus the average return of large cap stocks
- β_{iv} is the value factor loading of asset i, i.e., the asset's exposure to the value risk
- HML is the *High Minus Low* factor, computed as the average return of a portfolio composed of stocks with the highest Book-to-Market ratio minus the stocks with the lowest B/M ratio
- ε_i is the random error representing the idiosyncratic (company-specific) risk of asset i.

The Fama–French model does not respect the ideal environment: it cannot be clearly said that the market factor, HML and SMB are uncorrelated.

Portfolio Factor Model

If we have a single factor f, then a portfolio of n assets has the following properties

$$\begin{aligned} r_p = \sum_{i=1}^{n} r_i x_i &= \sum_{i=1}^{n} x_i(a_i + b_i f + \varepsilon_i) \\ &= \sum_{i=1}^{n} x_i a_i + (\sum_{i=1}^{n} x_i b_i) f + \sum_{i=1}^{n} x_i \varepsilon_i \\ &= a_p + b_p f + \varepsilon_p \end{aligned}$$

$$a_p = \sum_{i=1}^{n} x_i a_i, \qquad b_p = \sum_{i=1}^{n} x_i b_i, \qquad \varepsilon_p = \sum_{i=1}^{n} x_i \varepsilon_i$$

Effect of diversification

Assume $\text{var}(\varepsilon_i) = \sigma^2_{\varepsilon_i} = s^2 \quad \text{for } i = 1, ..., n.$

Let $x_i = 1/n$

$\sigma^2_{\varepsilon_p} = \text{var}(\sum_{i=1}^{n} x_i \varepsilon_i) = \sum_{i=1}^{n} x_i^2 \sigma^2_{\varepsilon_i} = s^2/n \to 0$ as $n \to \infty$

i.e., well diversified means $\quad \sigma^2_{\varepsilon_p} \approx 0 \quad (\varepsilon_p \approx 0)$

So $r_p = \alpha_p + \beta_p f$ for a well diversified portfolio.

2.5 Arbitrage Pricing Theory (APT)

Consider 2 well diversified funds, i and j:

Fund	Weight	Return
i	x	$r_i = a_i + b_i f$
j	$1 - x$	$r_j = a_j + b_j f$

$$r_p = x r_i + (1-x) r_j = \underbrace{x a_i + (1-x) a_j}_{a_p} + \underbrace{(x b_i + (1-x) b_j)}_{b_p} f$$

To make portfolio P riskless, choose x such that $b_p = 0$,
i.e., solve for x in $b_p = x b_i + (1-x) b_j = 0 \quad \to \quad x = \frac{b_j}{b_j - b_i}$

Plugging in we get:

$$r_p = \frac{a_i b_j}{b_j - b_i} + \frac{a_j b_i}{b_i - b_j} = r_f \quad \text{is the risk free rate}$$

$$r_f(b_j - b_i) = a_i b_j - a_j b_i$$

which can be rearranged to

$$\frac{a_j - r_f}{b_j} = \frac{a_i - r_f}{b_i} = c \quad \text{some constant (since it must hold for all } i \text{ and } j)$$

Through this relationship we can conclude that a_i and b_i are not independent.

$$a_i = r_f + b_i c$$

The takeaway from this derivation is to show that the values of a_i and b_i are not independent.

We can now show that

$$\begin{aligned} r_i &= a_i + b_i f \\ &= R_f + b_i c + b_i f \\ &= R_f + b_i(c + f) \end{aligned}$$

and

$$\begin{aligned} \mu_i &= a_i + b_i \bar{f} \\ &= r_f + b_i(c + \bar{f}) \\ &= \lambda_0 + b_i \lambda_1 \end{aligned}$$

where $\lambda_0 = r_f$ is the risk-free rate and $\lambda_1 = c + \bar{f}$ is considered the price of risk.

Example: An arbitrage opportunity

We live in a universe where the risk-free rate and the exact price of risk is known to be $\lambda_0 = 2\%$ and $\lambda_1 = 3\%$. Portfolios A and B have the factor loadings $b_A = 0.5$ and $b_B = 1$, respectively. Their expected return is

$$\begin{aligned} \mu_A &= \lambda_0 + b_A \lambda_1 = 0.02 + 0.5(0.03) = 3.5\% \\ \mu_B &= \lambda_0 + b_B \lambda_1 = 0.02 + 1(0.03) = 5\% \end{aligned}$$

Now consider a portfolio C with the following characteristics

$$\mu_C = 7.5\%, \quad b_C = 1.5$$

We can check for an arbitrage opportunity as follows

$$\lambda_0 + b_C \lambda_1 = 0.02 + 1.5(0.03) = 6.5\% \neq 7.5\%$$

This suggests that portfolio C is undervalued, since its expected return is higher than its sensitivity to the factor suggests. Therefore, we have found that an arbitrage opportunity is available. How do we exploit it?

→ We must construct a new portfolio by combining portfolios A and B that has the same factor loadings as C, let's call this new portfolio D. Once we find portfolio D, all we need to do is short portfolio D and buy an equal amount of C.

$$b_C = xb_A + (1-x)b_B \quad \Rightarrow \quad x = \frac{b_C - b_B}{b_A - b_B} = \frac{1.5-1}{0.5-1} = -1$$

If $x = -1$, then the expected return of portfolio D is

$$\mu_D = -\mu_A + 2\mu_B = -3.5 + 2(5) = 6.5\%.$$

We can short sell portfolio D and immediately use the funds to purchase portfolio C. By design, any risk we incur from this purchase is offset by our short sale of portfolio D since $b_C = b_D$.

General APT

Suppose there are n funds such that

$$r_i = a_i + \sum_{j=1}^{m} b_{ij} f_j, \quad m < n, \quad n \text{ funds governed by } m \text{ factors.}$$

Then there are constants $\lambda_0,\ \lambda_1,\ ...,\ \lambda_m$ such that

$$\mu_i = \lambda_0 + \sum_{j=1}^{m} b_{ij}\lambda_j \quad \text{for } i = 1, ..., n$$

In this case, the value of λ_j is the price of risk associated with the factor f_j.

2.6 Summary of MVO

Consider this version of MVO which minimizes risk subject to a target return constraint

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & x^T Q\, x \\ \text{s.t.} \quad & \mu^T x \geq R \\ & 1^T x = 1 \\ & (x_i \geq 0, \quad i = 1,\ ...,\ n) \end{aligned}$$

Would this problem be a convex problem?

Let's evaluate the objective function

$$\begin{aligned} f(x) &= x^T Q\, x \\ \nabla f(x) &= 2Qx \\ H(x) &= 2Q \end{aligned}$$

for any value of $x \in \mathbb{R}^n$. By construction, a covariance matrix is always symmetric positive semi-definite (SPSD), i.e., $u^T Q\, u \geq 0$ for any vector u. In practice, we may run into problems estimating our covariance matrix if we do not have enough observations. However, generally speaking, our square covariance matrix Q is SPSD.

$\rightarrow$ Since $H(x) = 2Q$, then the Hessian matrix is SPSD. Therefore, our objective function is convex over $x \in \mathbb{R}^n$. Any local min of a convex function is a global min, provided our solution space S is also convex.

Check the solution space (i.e., check if the constraints are convex)

By inspection of the coefficients in the constraints, we can see that they are all linear constraints (e.g,. $\mu^T x$ is linear). Luckily for us, all linear constraints are, by definition, convex. Therefore, our solution space S is convex.

Thus, our basic MVO problem proves to be a convex problem, meaning that any optimal solution we find will be a global optimal solution.

Methods to check if a point is optimal

If we do not have inequality constraints, then we can not only check if a point is optimal, we can find a closed-form solution to our problem by using the Lagrangian method. Consider the *Minimum Variance* portfolio with short-selling allowed

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \frac{1}{2} x^T Q \; x \\ \text{s.t.} \quad & \mathbf{1}^T x = 1 \end{aligned}$$

This simplified version of MVO has only one equality constraint, the budget constraint.

Example:

$$Q = \begin{bmatrix} 0.8 & 0.05 & -0.1 \\ 0.05 & 0.6 & 0.08 \\ -0.1 & 0.08 & 0.7 \end{bmatrix}, \qquad c = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad A = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}, \qquad b = 1$$

Recall the Lagrangian formulation of a generic quadratic program

$$\begin{aligned} \mathcal{L} &= \frac{1}{2} x^T Q \; x + c^T x + \pi(Ax - b) \\ \frac{\partial \mathcal{L}}{\partial x} &= Q \; x + c + A^T \pi = 0 \\ \frac{\partial \mathcal{L}}{\partial \pi} &= Ax - b = 0 \end{aligned}$$

In the context of MVO, the Lagrangian is

$$\begin{aligned} \mathcal{L} &= \frac{1}{2} x^T Q \; x + \pi(\mathbf{1}^T x - 1) \\ \frac{\partial \mathcal{L}}{\partial x} &= Q \; x + \mathbf{1}\pi = 0 \\ \frac{\partial \mathcal{L}}{\partial \pi} &= \mathbf{1}^T x - 1 = 0 \end{aligned}$$

where π is the corresponding Lagrangian multiplier. Now we can find a closed-form solution

$$x = -Q^{-1}(\mathbf{1}\pi) \quad \text{plug into second equation}$$
$$\mathbf{1}^T Q^{-1}(\mathbf{1}\pi) = -1 \quad \rightarrow \quad \text{solve for } \pi, \quad \pi^* = -(\mathbf{1}^T Q^{-1}\mathbf{1})^{-1}1$$
$$\text{Solving for } x \quad \rightarrow \quad x^* = Q^{-1}(\mathbf{1}(\mathbf{1}^T Q^{-1}\mathbf{1})^{-1}b)$$

Plugging in our values,

$$Q^{-1} = \begin{bmatrix} 1.28 & -0.13 & 0.20 \\ -0.13 & 1.71 & -0.21 \\ 0.20 & -0.21 & 1.48 \end{bmatrix}$$

$$\pi^* = -0.24, \qquad x^* = \begin{bmatrix} 0.323 \\ 0.326 \\ 0.351 \end{bmatrix}$$

Checking for optimality (cont'd): KKT Conditions

What if we want to check if a portfolio $\bar{x}$ is the *Minimum Variance* portfolio with no short-selling allowed?

$$\begin{aligned} \min_{x} \quad & \frac{1}{2} x^T Q\, x \\ \text{s.t.} \quad & \mathbf{1}^T x = 1 \\ & x \geq 0 \end{aligned}$$

Let

$$\bar{x} = \begin{bmatrix} 0.323 \\ 0.326 \\ 0.351 \end{bmatrix}$$

Check feasibility:

$$h(\bar{x}) = 0.323 + 0.326 + 0.351 - 1 = 0 \quad \therefore \text{ feasible}$$
$$g_1(\bar{x}) = -0.323 < 0 \quad \therefore \text{ inactive}$$
$$g_2(\bar{x}) = -0.326 < 0 \quad \therefore \text{ inactive}$$
$$g_3(\bar{x}) = -0.351 < 0 \quad \therefore \text{ inactive}$$

Since all inequality constraints are inactive, we must have that $u_1 = u_2 = u_3 = 0$ to guarantee complementary slackness (which is the 3rd KKT condition).

Setup our KKT conditions:

1. Stationarity: $Q\ \bar{x} - I\ u + A^T v = 0$

$$\begin{bmatrix} 0.8 & 0.05 & -0.1 \\ 0.05 & 0.6 & 0.08 \\ -0.1 & 0.08 & 0.7 \end{bmatrix} \begin{bmatrix} 0.323 \\ 0.326 \\ 0.351 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} v = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

Since $u_1 = u_2 = u_3 = 0$, we have

$$\begin{bmatrix} 0.240 \\ 0.240 \\ 0.240 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} v = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \rightarrow \quad v = -0.24$$

2. Dual feasibility: $u = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \geq 0$
3. Complementary slackness: $u_i \cdot g_i(x) = 0 \cdot (-x_i) = 0$ for $i = 1, 2, 3$

This means that $\bar{x}$ complies with the KKT conditions and may be a local min. We can check if this problem complies with the SONC at the point $\bar{x}$. In our case, we can show that Q is PSD/PD by checking its principal minors or its eigenvalues.

$$\Delta_1 = 0.8, \quad \Delta_2 = 0.477, \quad \Delta_3 = 0.322 \quad \rightarrow \quad Q \text{ is PD, independent of } x$$

Since our constraints are linear, we can conclude that the feasible set is also convex for all $x \in \mathbb{R}^3$, meaning the SONC becomes sufficient. Moreover, since Q is PD we can also say our optimal solution is unique. Thus, $\bar{x}$ is a <u>unique global minimum</u>.