Portfolio Optimization # 2
A. Charpentier (Université de Rennes 1)
Markowitz (1952) & Theoretical Approach

Following Markowitz (1952), consider $n$ assets, infinitely divisible. Their returns are random variables, denoted $X$, (jointly) normally distributed, $\mathcal{N}(\mu, \Sigma)$, i.e. $E[X] = \mu$ and $\text{var}(X) = \Sigma$.

Let $\omega$ denotes weights of a given portfolio.

Portfolio risk is measured by its variance $\text{var}(\alpha^T X) = \sigma^2_\alpha = \alpha^T \Sigma \alpha$

For minimal variance portfolio, the optimization problem can be stated as

$$\alpha^* = \arg\min\{\alpha^T \Sigma \alpha\} \text{ s.t. } \alpha^T 1 \leq 1$$
No short sales

One can add a no short sales contraints, i.e. $\alpha \in \mathbb{R}^n_+$. Thus, we should solve

$$\alpha^* = \arg\min \{ \alpha^T \Sigma \alpha \} \text{ s.t. } \begin{cases} -\alpha^T 1 \geq -1 \\ \alpha_i \geq 0, \forall i = 1, \ldots, n, \end{cases}$$

The later constraints can also be written

$$A^T \alpha \geq b$$

where

$$A^T = \begin{pmatrix} -1 & -1 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$
> asset.names = c("MSFT", "NORD", "SBUX")
> muvec = c(0.0427, 0.0015, 0.0285)
> names(muvec) = asset.names
> sigmamat = matrix(c(0.0100, 0.0018, 0.0011, 0.0018, 0.0109, 0.0026,
                     0.0011, 0.0026, 0.0199), nrow=3, ncol=3)
> dimnames(sigmamat) = list(asset.names, asset.names)
> r.f = 0.005
> cov2cor(sigmamat)

<table>
<thead>
<tr>
<th></th>
<th>MSFT</th>
<th>NORD</th>
<th>SBUX</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSFT</td>
<td>1.000</td>
<td>0.172</td>
<td>0.078</td>
</tr>
<tr>
<td>NORD</td>
<td>0.172</td>
<td>1.000</td>
<td>0.177</td>
</tr>
<tr>
<td>SBUX</td>
<td>0.078</td>
<td>0.177</td>
<td>1.000</td>
</tr>
</tbody>
</table>
No short sales

One can use the `solve.QP` function of `library(quadprog)`, which solves

\[
\min_{\mathbf{x} \in \mathbb{R}^n} \left\{ \frac{1}{2} \mathbf{x}^T \mathbf{D} \mathbf{x} - \mathbf{d}^T \mathbf{x} \right\} \quad \text{s.t.} \quad \mathbf{A}^T \mathbf{x} \geq \mathbf{b}
\]

```r
1 > Dmat = 2*sigmamat
2 > dvec = rep(0,3)
3 > Amat = cbind(rep(1,3),diag(3))
4 > bvec = c(1,0,0,0)
5 > opt = solve.QP(Dmat,dvec,Amat,bvec,meq=1)
6 > opt$solution
7 [1] 0.441 0.366 0.193
```
No short sales

One can also use

```r
> gmin.port = globalMin.portfolio(muvec, sigmamat, shorts=FALSE)
> gmin.port

Call:
globalMin.portfolio(er = mu.vec, cov.mat = sigma.mat, shorts = FALSE)

Portfolio expected return: 0.0249
Portfolio standard deviation: 0.0727
Portfolio weights:
  MSFT  NORD  SBUX
0.441  0.366  0.193
```
No short sales

Consider minimum variance portfolio with same mean as Microsoft (see #1),

\[
\alpha^* = \arg \min \{ \alpha^T \Sigma \alpha \} \quad \text{s.t.} \quad \begin{cases}
\mathbb{E}(\alpha^T X) = \alpha^T \mu \geq \bar{r} = \mu_1 \\
\alpha^T 1 \leq 1
\end{cases}
\]

> (eMsft.port = efficient.portfolio(mu.vec, sigma.mat, target.return = mu.vec["MSFT"]))

<table>
<thead>
<tr>
<th>Portfolio expected return:</th>
<th>0.0427</th>
</tr>
</thead>
<tbody>
<tr>
<td>Portfolio standard deviation:</td>
<td>0.0917</td>
</tr>
<tr>
<td>Portfolio weights:</td>
<td></td>
</tr>
<tr>
<td>MSFT</td>
<td>NORD</td>
</tr>
<tr>
<td>0.8275</td>
<td>-0.0907</td>
</tr>
</tbody>
</table>

There is some short sale here.
No short sales

\[ \alpha^* = \arg\min \{ \alpha^T \Sigma \alpha \} \quad \text{s.t.} \quad \begin{align*}
\alpha^T \mu & \geq \bar{r} \\
-\alpha^T \mathbf{1} & \geq -1 \\
\alpha_i & \geq 0, \forall i = 1, \cdots, n,
\end{align*} \]

The later constraints can also be written

\[ A^T \alpha \geq b \]

where

\[ A^T = \begin{pmatrix}
\mu_1 & \mu_2 & \mu_3 \\
-1 & -1 & -1 \\
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{pmatrix}, \quad \text{and} \quad b = \begin{pmatrix}
\bar{r} \\
-1 \\
0 \\
0 \\
0
\end{pmatrix} \]
No short sales

```r
> Dmat = 2*sigma.mat
> dvec = rep(0, 3)
> Amat = cbind(mu.vec, rep(1,3), diag(3))
> bvec = c(mu.vec["MSFT"], 1, rep(0,3))
> qp.out = solve.QP(Dmat=Dmat, dvec=dvec, Amat=Amat, bvec=bvec, meq=2)
> names(qp.out$solution) = names(mu.vec)
> round(qp.out$solution, digits=3)
```

<table>
<thead>
<tr>
<th></th>
<th>MSFT</th>
<th>NORD</th>
<th>SBUX</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
No short sales

We can also get the efficient frontier, if we allow for short sales,

```r
> ef = efficient.frontier(mu.vec, sigma.mat, alpha.min=0, alpha.max=1, nport=10)
> ef$weights
    MSFT  NORD  SBUX
port 1  0.827 -0.0907  0.263
port 2  0.785 -0.0400  0.256
port 3  0.742  0.0107  0.248
port 4  0.699  0.0614  0.240
port 5  0.656  0.1121  0.232
port 6  0.613  0.1628  0.224
port 7  0.570  0.2135  0.217
port 8  0.527  0.2642  0.209
port 9  0.484  0.3149  0.201
port10 0.441  0.3656  0.193
```
No short sales
No short sales

It is easy to compute the efficient frontier without short sales by running a simple loop

```r
mu.vals = seq(gmin.port$er, max(mu.vec), length.out=10)
w.mat = matrix(0, length(mu.vals), 3)
sd.vals = rep(0, length(sd.vec))
colnames(w.mat) = names(mu.vec)
D.mat = 2*sigma.mat
d.vec = rep(0, 3)
A.mat = cbind(mu.vec, rep(1,3), diag(3))
for (i in 1:length(mu.vals)) {
  b.vec = c(mu.vals[i],1, rep(0,3))
  qp.out = solve.QP(Dmat=D.mat, dvec=d.vec,
                    Amat=A.mat, bvec=b.vec, meq=2)
  w.mat[i,] = qp.out$solution
  sd.vals[i] = sqrt(qp.out$value)
}
```
No short sales
```r
> ef.ns = efficient.frontier(mu.vec, sigma.mat, alpha.min=0, alpha.max=1, nport=10, shorts=FALSE)

ef.ns$weights

<table>
<thead>
<tr>
<th></th>
<th>MSFT</th>
<th>NORD</th>
<th>SBUX</th>
</tr>
</thead>
<tbody>
<tr>
<td>port 1</td>
<td>0.441</td>
<td>0.3656</td>
<td>0.193</td>
</tr>
<tr>
<td>port 2</td>
<td>0.484</td>
<td>0.3149</td>
<td>0.201</td>
</tr>
<tr>
<td>port 3</td>
<td>0.527</td>
<td>0.2642</td>
<td>0.209</td>
</tr>
<tr>
<td>port 4</td>
<td>0.570</td>
<td>0.2135</td>
<td>0.217</td>
</tr>
<tr>
<td>port 5</td>
<td>0.613</td>
<td>0.1628</td>
<td>0.224</td>
</tr>
<tr>
<td>port 6</td>
<td>0.656</td>
<td>0.1121</td>
<td>0.232</td>
</tr>
<tr>
<td>port 7</td>
<td>0.699</td>
<td>0.0614</td>
<td>0.240</td>
</tr>
<tr>
<td>port 8</td>
<td>0.742</td>
<td>0.0107</td>
<td>0.248</td>
</tr>
<tr>
<td>port 9</td>
<td>0.861</td>
<td>0.0000</td>
<td>0.139</td>
</tr>
<tr>
<td>port 10</td>
<td>1.000</td>
<td>0.0000</td>
<td>0.000</td>
</tr>
</tbody>
</table>
```
Independent observations?

Let \( U = (U_1, \ldots, U_n)^T \) denote a random vector such that \( U_i \)'s are uniform on \([0, 1]\), then its cumulative distribution function \( C \) is called a copula, defined as

\[
C(u_1, \cdots, u_n) = \Pr[U \leq u] = \mathbb{P}[U_1 \leq u_1, \cdots, U_n \leq u_n], \forall u \in [0, 1]^n.
\]

Let \( Y = (Y_1, \ldots, Y_n)^T \) denote a random vector with cdf \( F(\cdot) \), such that \( Y_i \) has marginal distribution \( F_i(\cdot) \).

From Sklar (1959), there exists a copula \( C : [0, 1]^d \to [0, 1] \) such that

\[
\forall y = (y_1, \ldots, y_n)^T \in \mathbb{R}^n,
\]

\[
\mathbb{P}[Y \leq y] = F(y) = C(F_1(y_1), \ldots, F_n(y_n)).
\]

Conversely, set \( U_i = F_i(Y_i) \). If \( F_i(\cdot) \) is absolutely continuous, \( U_i \) is uniform on \([0, 1]\), \( C(\cdot) \) is the cumulative distribution function of \( U = (U_1, \ldots, U_n)^T \). Set \( F_i^{-1}(u) = \inf\{x_i, F_i(x_i) \geq u\} \) then

\[
C(u_1, \cdots, u_n) = \mathbb{P}[Y_1 \leq F_1^{-1}(u_1), \cdots, Y_n \leq F_n^{-1}(u_n)], \forall u \in [0, 1]^n.
\]
Independent observations?

The Gaussian copula with correlation matrix $R$ is defined as

$$C(u|R) = \Phi_R(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_n)) = \Phi_R(\Phi^{-1}(u)),$$

where $\Phi_R$ is the cdf of $\mathcal{N}(0, R)$. Thus copula has density

$$c(u|R) = \frac{1}{\sqrt{|R|}} \exp \left( -\frac{1}{2} \Phi^{-1}(u)^T [R^{-1} - \mathbb{I}] \Phi^{-1}(u) \right).$$
Independent observations?

The Student \( t \)-copula, with correlation matrix \( R \) and with \( \nu \) degrees of freedom has density

\[
c(u|R, \nu) = \frac{\Gamma \left( \frac{\nu+n}{2} \right) \Gamma \left( \frac{\nu}{2} \right)^n \left[ 1 + \frac{1}{\nu} T_{\nu}^{-1}(u)^\top R^{-1} T_{\nu}^{-1}(u) \right]^{-\frac{\nu+n}{2}}}{|R|^{1/2} \Gamma \left( \frac{\nu+n}{2} \right) \Gamma \left( \frac{\nu}{2} \right) \prod_{i=1}^n \left[ 1 + \frac{T_{\nu}(u_i)^2}{\nu} \right]^{-\frac{\nu+1}{2}}}
\]

where \( T_{\nu} \) is the cdf of the Student-\( t \) distribution.

One can also consider, instead of the induced copula of the Student-\( t \) distribution, the one associated with the \( s \) skew-\( t \) (see Genton (2004))

\[
x \mapsto t(x|\nu, R) \cdot T_{\nu+d} \left( s^\top R^{-1/2} x \sqrt{\frac{\nu + d}{x^\top R^{-1} x + \nu}} \right)
\]
Independent Return?

It might be more realistic to consider the conditional distribution of $Y_t$ given information $\mathcal{F}_{t-1}$.

Assume that $Y_t$ has conditional distribution $F(\cdot|\mathcal{F}_{t-1})$, such that $\forall i$, $Y_{i,t}$ has conditional distribution $F_i(\cdot|\mathcal{F}_{t-1})$.

There exists $C(\cdot|\mathcal{F}_{t-1}) : [0,1]^d \to [0,1]$ such that $\forall y = (y_1,\ldots,y_n)^T \in \mathbb{R}^n$,

$$
P[Y_t \leq y|\mathcal{F}_{t-1}] = F(y|\mathcal{F}_{t-1}) = C(F_1(y_1|\mathcal{F}_{t-1}),\ldots,F_n(y_n|\mathcal{F}_{t-1})|\mathcal{F}_{t-1}).$$

Conversely, set $U_{i,t} = F_i(Y_{i,t}|\mathcal{F}_{t-1})$. If $F_i(\cdot|\mathcal{F}_{t-1})$ is absolutely continuous, $C(\cdot|\mathcal{F}_{t-1})$ is the cdf of $U_t = (U_{1,t},\ldots,U_{d,t})^T$, given $\mathcal{F}_{t-1}$.

Consider weekly returns of oil prices, Brent, Dubaï and Maya.
One can consider $Y_t = \mu_t + \Sigma_t^{1/2} \varepsilon_t$, where $\Sigma_t = \tilde{\Sigma}^{1/2} R_t \Sigma_t^{1/2}$, where $\tilde{\Sigma}$ is a diagonal matrix (with only variance terms).

More generally $\mathbb{E}(Y_{i,t} | \mathcal{F}_{t-1}) = \mu_i(Z_{t-1}, \alpha)$ and $\text{var}(Y_{i,t} | \mathcal{F}_{t-1}) = \sigma_i^2(Z_{t-1}, \alpha)$ for some $Z_{t-1} \in \mathcal{F}_{t-1}$.

Those include VAR models (with conditional means) and MGARCH models (with conditional variances).
VAR models

See Sims (1980) and Lütkepohl (2007)

\[
E(Y_{i,t}|\mathcal{F}_{t-1}) = \mu_{i,0} + \mu_{i}(Y_{t-1}, \Phi, \Sigma) = \Phi_i^T Y_{t-1},
\]

and

\[
\text{var}(Y_{i,t}|\mathcal{F}_{t-1}) = \sigma_{i}^2(Y_{t-1}, \Phi, \Sigma) = \Sigma_{i,i}.
\]

Or fit univariate models,

```r
> fit1 = arima(x=dat[,1], order=c(2,0,1))
> library(fGarch)
> fit1 = garchFit(formula = ~ arma(2,1)+garch(1, 1), data=dat[,1], cond
  .dist ="std", trace=FALSE)
```
MGARCH models

See Engle & Kroner (1995) for the BEEK model

Consider $Y_t = \Sigma_t^{1/2} \varepsilon_t$ where $\mathbb{E}(\varepsilon_t) = 0$ and $\text{var}(\varepsilon_t) = I$.

The MGARCH(1,1) is defined as

$$\Sigma_t = \omega^T \omega + A^T \varepsilon_{t-1} \varepsilon_{t-1}^T A + B^T \Sigma_{t-1} B$$

where $\omega$, $A$ et $B$ are $d \times d$ matrices ($\omega$ can be triangular matrix). Then

$$\mathbb{E}(Y_{i,t} | \mathcal{F}_{t-1}) = \mu_i(\Sigma_{t-1}, \varepsilon_{t-1}, \omega, A, B, \Sigma) = 0$$

and $\Sigma_t = \mathbb{V}(Y_t | \mathcal{F}_{t-1})$ with

$$\text{var}(Y_{i,t} | \mathcal{F}_{t-1}) = \sigma_i^2(\Sigma_{t-1}, \varepsilon_{t-1}, \omega, A, B, \Sigma) = \Sigma_{i,i,t}$$

i.e.

$$\text{var}(Y_{i,t} | \mathcal{F}_{t-1}) = \omega_i^T \omega_i + A_i^T \varepsilon_{t-1} \varepsilon_{t-1}^T A_i + B_i^T \Sigma_{t-1} B_i.$$
MGARCH models: BEKK, $(\Sigma_t)$

1. `library(MTS)`
2. `bekk=BEKK11(data_res)`

Evolution of $t \mapsto (\Sigma_{i,i,t})$.
MGARCH models: BEKK, \((R_t)\)

Evolution of \(t \mapsto (R_{i,i,t})\).
Copula based models

Here $Y_{i,t} = \mu_i(Z_{t-1}, \alpha) + \sigma_i(Z_{t-1}, \alpha) \cdot \varepsilon_{i,t}$, thus, define standardized residuals

$$
\varepsilon_{i,t} = \frac{Y_{i,t} - \mu_i(Z_{t-1}, \alpha)}{\sigma_i(Z_{t-1}, \alpha)}.
$$

Recall that the (conditional) correlation matrix is $R_t = \tilde{\Sigma}_t^{-1/2} \Sigma_t \tilde{\Sigma}_t^{-1/2}$, where $\tilde{\Sigma}_t = \text{diag}(\Sigma_t)$, with

$$
R_{i,j,t} = \text{cor}(\varepsilon_{i,t}, \varepsilon_{j,t}|F_{t-1}) = E[\varepsilon_{i,t}\varepsilon_{j,t}|F_{t-1}].
$$

Engle (2002) suggested $R_t = \tilde{Q}_t^{-1/2} Q_t \tilde{Q}_t^{-1/2}$ where $Q_t$ is driven by

$$
Q_t = [1 - \alpha - \beta] \bar{Q} + \alpha Q_{t-1} + \beta \varepsilon_{t-1} \varepsilon_{t-1}^T,
$$

where $\bar{Q}$ is the unconditional variance of $(\eta_t)$, and where coefficients $\alpha$ and $\beta$ are positive, with $\alpha + \beta \in (0, 1)$. 
Time varying correlation ($R_t$)
EWMA : Exponentially Weighted Moving Average

In the univariate case,

$$\sigma_t^2 = (1 - \lambda)[r_{t-1} - \mu_{t-1}]^2 + \lambda \sigma_{t-1}^2,$$

with $\lambda \in [0, 1)$. It is called **Exponentially Weighted Moving Average** because

$$\sigma_t^2 = [1 - \lambda] \sum_{i=1}^{n} \lambda^i \sigma_{t-i}^2 + \lambda^n \sigma_{t-n}^2 = [1 - \lambda] \sum_{i=1}^{\infty} \lambda^i \sigma_{t-i}^2.$$

The (natural) multivariate extension would be

$$\Sigma_t = (1 - \lambda)(r_{t-1} - \mu_{t-1})(r_{t-1} - \mu_{t-1})^T + \Lambda \Sigma_{t-1},$$

in the sense that

$$\Sigma_{i,j,t} = (1 - \lambda)[r_{i,t-1} - \mu_{i,t-1}][r_{j,t-1} - \mu_{j,t-1}] + \lambda \Sigma_{i,j,t-1}$$

see Lowry et al. (1992).
EWMA : Exponentially Weighted Moving Average

Best (1998) suggested

\[
\Sigma_{i,j,t} \approx [1 - \lambda] \sum_{k=1}^{t} \lambda^i [r_{i,t-k} - \mu_{i,t-k}] [r_{j,t-k} - \mu_{j,t-k}]
\]

i.e.

\[
\Sigma_t = [1 - \lambda] \sum_{k=1}^{t} \lambda^k (r_{t-k} - \mu_{t-k})(r_{t-k} - \mu_{t-k})^T + \Lambda^t \Sigma_0.
\]

(here \(\Sigma_t\) is a definite-positive matrix when \(\Sigma_0\) is a variance matrix).

More generally

\[
\Sigma_t = (1 - \Lambda_1)(r_{t-1} - \mu_t)(r_{t-1} - \mu_t)^T (1 - \Lambda_1)^T + \Lambda_1 \Sigma_{t-1} \Lambda_1^T,
\]

where \(\mu_t = (1 - \lambda_2)r_{t-1} + \lambda_2 \mu_{t-1}\), \(\Lambda_1\) being a diagonal matrix, \(\Lambda_1 = \text{diag}(\lambda_1)\), and \(\lambda_2\) being vectors of \(\mathbb{R}^n\).
**EWMA :** $(\Sigma_t)$

1. `library(MTS)`
2. `ewma=EWMAvol(dat_res, lambda = 0.96)`

Evolution of $t \mapsto (\Sigma_{i,i,t})$ for the three oil series.
EWMA: \((R_t)\)

From matrices \(\Sigma_t\), one can extract correlation matrices \(R_t\), since
\[
R_t = \tilde{\Sigma}_t^{-1/2} \Sigma_t \tilde{\Sigma}_t^{-1/2}
\]

Evolution de \((R_{i,j,t})\) for the three pairs of oil series.
When Variance and Returns are Time Varying

Consider the case where means and variances are estimated at time $t$ based on $\{y_{t-60}, \cdots, y_{t-1}, y_t\}$, and let $\hat{\alpha}_t^*$ denote the allocation of the minimal variance portfolio, when there are no short sales.
When Variance and Returns are Time Varying

Consider the value of the portfolio either with fixed weights $\tilde{\alpha}_{60}^*$, or with weights updated every month, $\tilde{\alpha}_t^*$. 
Moving Away from the Variance

Consider various risk measures, on losses $X$

- Value-at-Risk, $\text{VaR}_X(\alpha) = F^{-1}(X)$
- expected shortfall, $\text{ES}^+_X(\alpha) = \mathbb{E}[X|X \geq \text{VaR}_X(\alpha)]$
- conditional VaR, $\text{CVaR}_X^+(\alpha) = \mathbb{E}[X - \text{VaR}_X(\alpha)|X \geq \text{VaR}_X(\alpha)]$

One can use `library(FRAPO)` and `library(fPortfolio)` to compute the minimum CVaR portfolio.

```r
> cvar = portfolioSpec()
> setType(cvar) = "CVaR"
> setAlpha(cvar) = 0.1
> setSolver(cvar) = "solveRglpk.CVAR"
> minriskPortfolio(data = X, spec = cvar, constraints = "LongOnly")
```
Moving Away from the Variance

Consider the case where means and CVaR_{10\%} are estimated at time \( t \) based on \( \{y_{t-60}, \cdots, y_{t-1}, y_t\} \), and let \( \hat{\alpha}_t^* \) denote the allocation of the minimal CVaR_{10\%} portfolio, when there are no short sales.
Moving Away from the Variance

Consider the value of the portfolio either with fixed weights $\alpha_6^*$, either for minimal variance, or minimal $\text{CVaR}_{10\%}$.
Moving Away from the Variance

Consider the value of the portfolio either with weights updated every month, $\hat{\alpha}_t^*$, either for minimal variance, or minimal CVaR$_{10\%}$. 

![Graph showing portfolio value over time](image-url)
Moving Away from the Variance

One can also consider optimal portfolios $\hat{\alpha}_t^\star$, for minimal CVaR$_{10\%}$, CVaR$_{5\%}$ and CVaR$_{1\%}$.