Dynamic Programming
動態規劃

張智星 (Roger Jang)

jang@mirlab.org

http://mirlab.org/jang

多媒體資訊檢索實驗室

台灣大學 資訊工程系

Dynamic Programming

- Dynamic Programming (DP)
  - An effective method for finding the optimum solution to a multi-stage decision problem, based on the principal of optimality
- Applications: NUMEROUS!

Quiz!

  - Longest common subsequence, edit distance, matrix chain products, all-pair shortest distance, dynamic time warping, hidden Markov models, …

Principal of Optimality

- Richard Bellman, 1952
  - An optimal policy has the property that whatever the initial state and the initial decisions are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.

Web Resources about DP

- Recordings on the web
  - [MIT Open Course Ware]

Problems Solvable by DP

- Characteristics of problems solvable by DP

  Quiz!

  - Decomposition: The original problem can be expressed in terms of subproblems.
  - Subproblem optimality: the global optimum value of a subproblem can be defined in terms of optimal subproblems of smaller sizes.

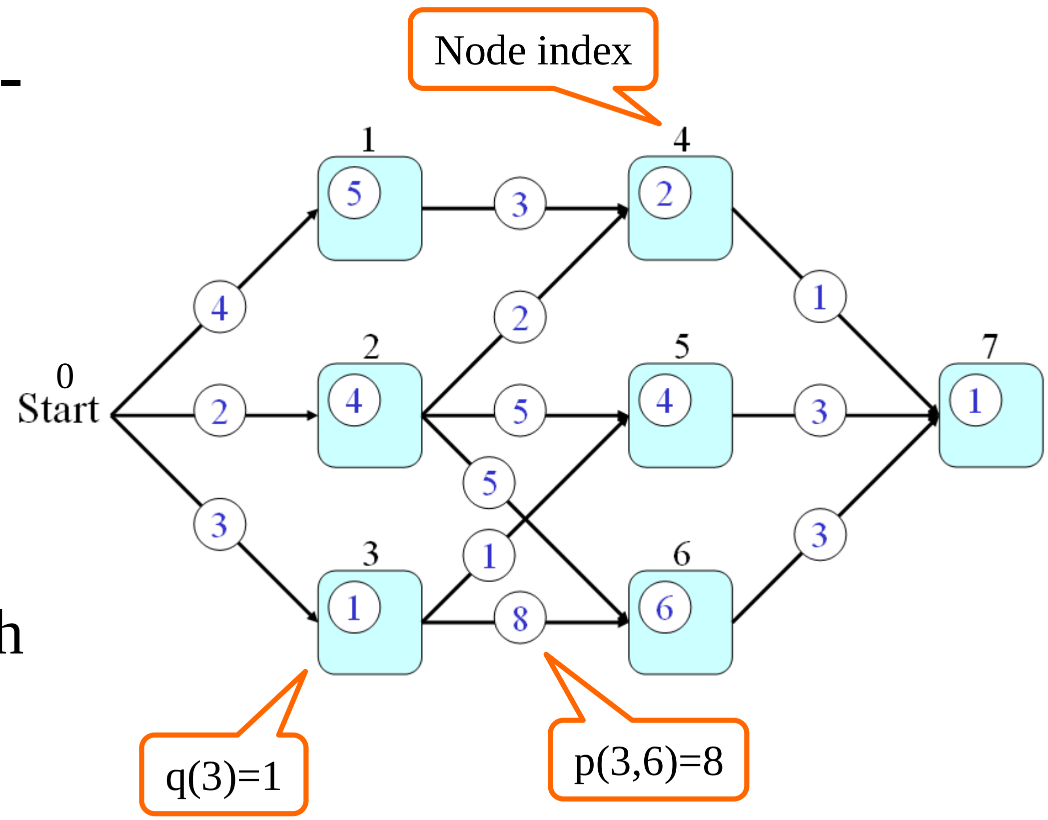
Three-step Formulas of DP

Quiz!

- DP formulation involves 3 steps
  - Define the optimum-value function for recursion
  - Derive the recurrent formula of the optimum-value function, with boundary conditions
  - Specify the answer to the original task in terms of the optimum-value function.

DP Example: Optimal Path Finding

- Path finding in a feed-forward network
  - p(a,b): transition cost
  - q(a): state cost
- Goal
  - Find the optimal path from nodes 0 to 7 such that the total cost is minimized.



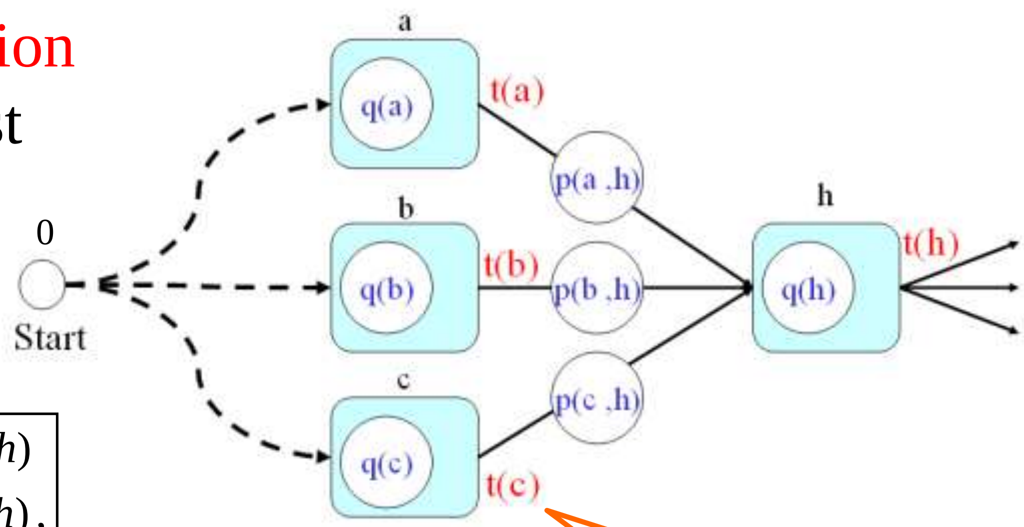
DP Example: Optimal Path Finding

- Three steps in DP
  - Optimum-value function
    t(h): the minimum cost from the start point (node 0) to node h.
  - Recurrent formula

$$t(h)=q(h)+\min\begin{cases} t(a)+p(a,h) \\ t(b)+p(b,h), \\ t(c)+p(c,h) \end{cases}$$

with boundary condition $t(0)=0$.

  - Answer: t(7)

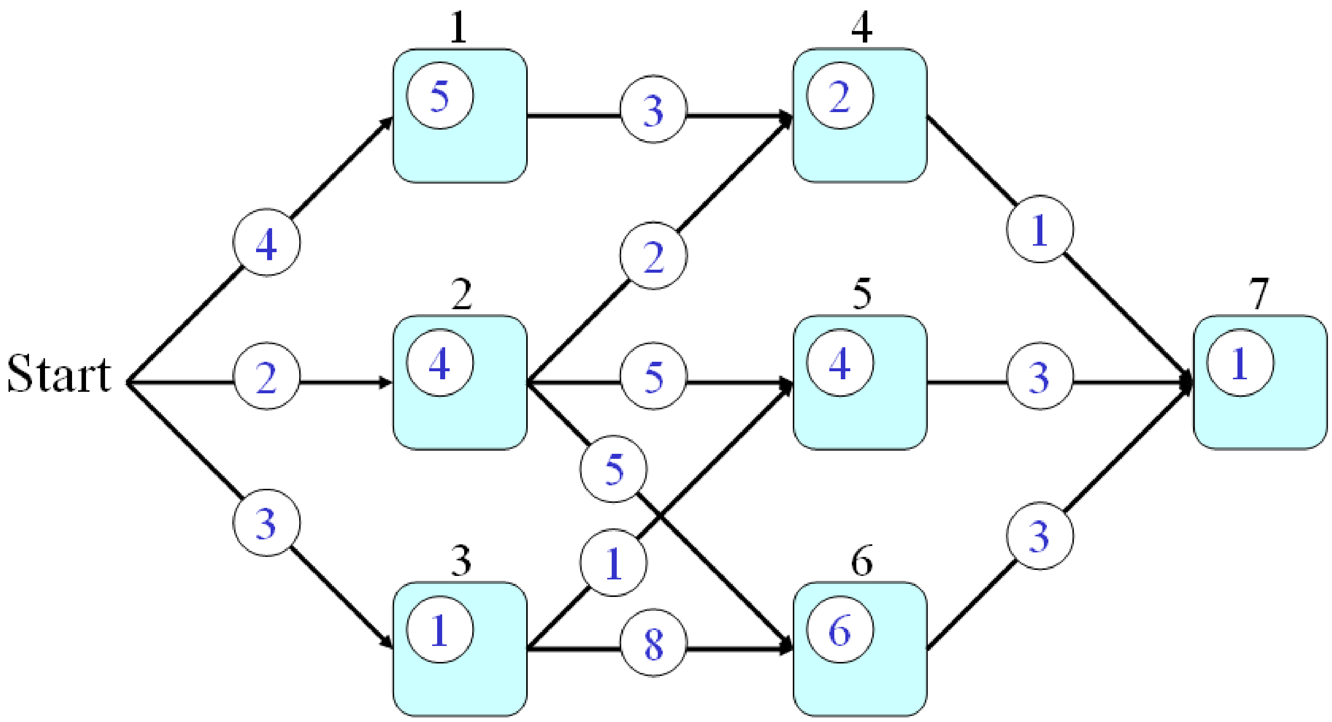


Optimum-value function

Last node

DP Example: Optimal Path Finding

- Walkthrough of DP
  - Flash
  - Gif

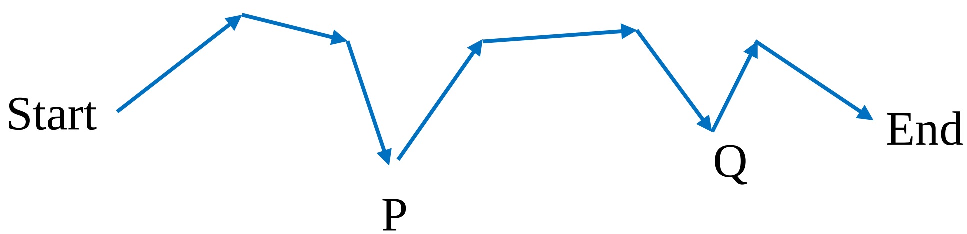


Observations

- Some observations based on this path finding example
  - Once t(7) is found, t(k), $\forall k<7$ is also found
  - Multi-stage → Layer-by-layer computation → No loops in the graph
- In fact
  - Any DP problem can be visualized as this optimal path finding problem!

Principal of Optimality: Example

- In terms of the min-cost path problem
  - Any sub-path of the min-cost path should itself be a min-cost path given the starting and ending nodes



P → Q
^Q → ^P

Bottom-up Approach of DP

- Usually bottom-up design of DP
  - Start from the bottom (base cases)
  - Solve small sub-problems
  - Store solutions
  - Reuse previous results for solving larger sub-problems

Usually it's reduced to path finding via table filling!

Characteristics of DP

- Some general characteristics of DP

Quiz!

  - We need to store back-tracking information in order to identify the path efficiently.
  - Once the optimal path is found, all the related sub-problems are also solved.
  - DP can only find the optimal path. To find the second best, we need to resort to a more complicated n-best approach.

Back to the Future: When to Buy and Sell?

J.-S. Roger Jang (張智星)
jang@mirlab.org
http://mirlab.org/jang
MIR Lab, CSIE Dept.
National Taiwan University

2023/10/3

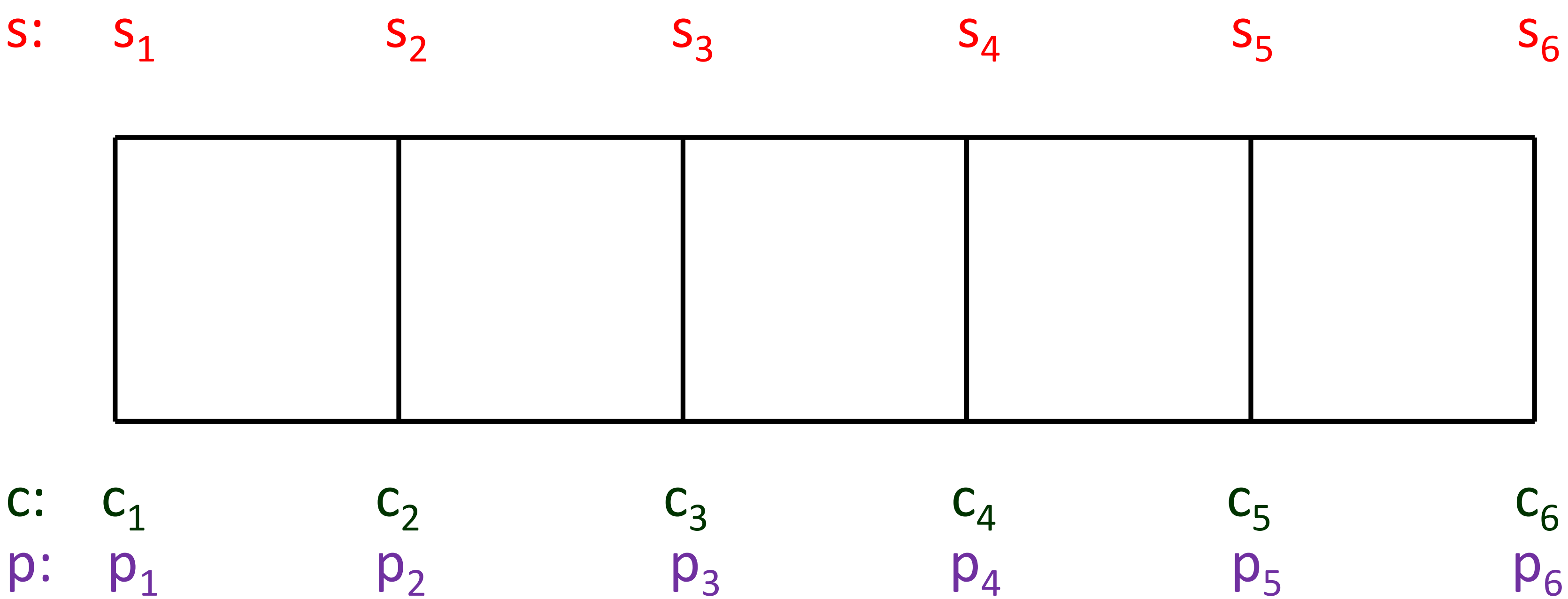
Introduction

- Why “Back to the Future”?
- Problem: Given the price of a stock over a span of time, how can you determine when to “buy” and “sell” to maximize the overall return?
  - Assumptions
    - Each day has a single price for a stock.
    - You can buy or sell only once in a day.
    - You can always get the transaction done.
    - Transaction fee applies.
    - Always “buy all” or “sell all”.
- Analytic solution exists ➔ DP!

DP Formula for Trading

Quiz!

- Notations
  - p_i: stock price at stage i
  - s_i: max. stockholding at stage i
  - c_i: max. cash at stage i, with c_1 being the initial cash
- Recurrent formula



Example of DP for Trading

Quiz!

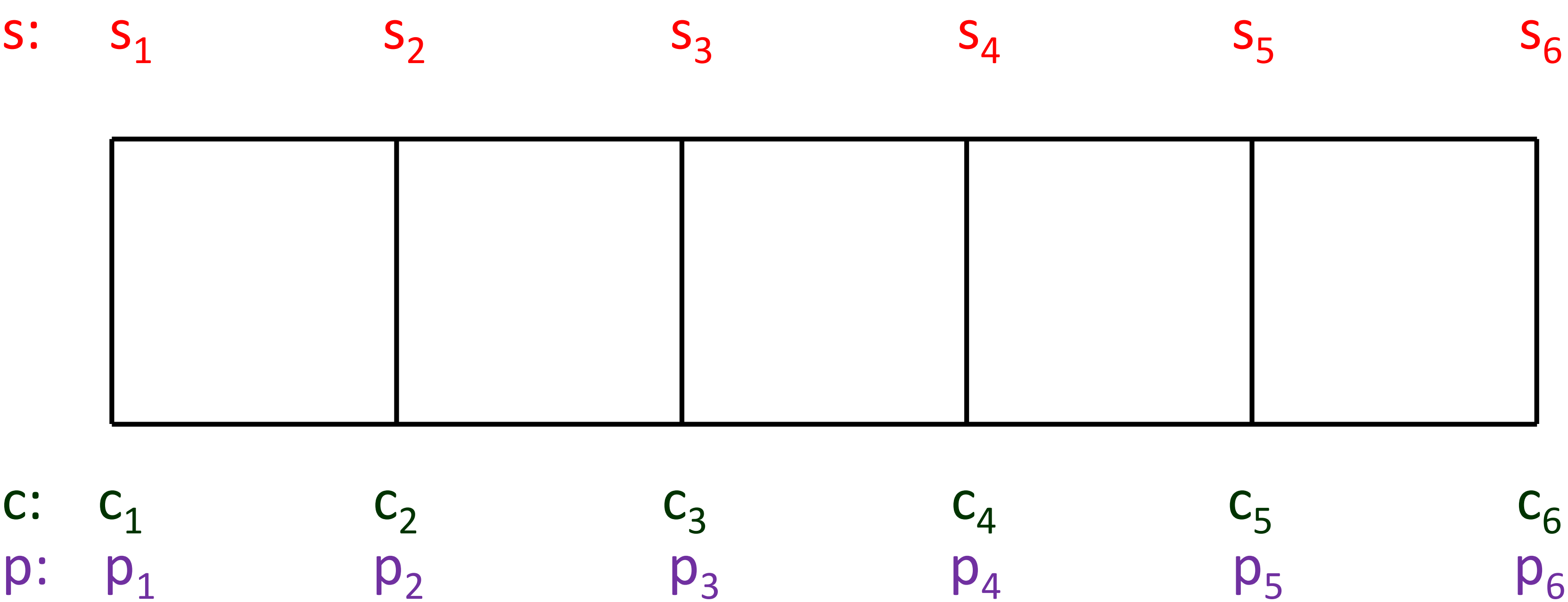
- Given p = [2 3 4 2 3 2], c_1=10
- Compute c_i and s_i, i=1~6
- Recurrent formula

s:	5	5	5	10	10	15
c:	10	15	20	20	30	30
p:	2	3	4	2	3	2

DP Formula for Trading, with Transaction Fee

Quiz!

- Notations: same as before
  - ρ: rate for transaction fee
- Recurrent formula



Extension

- Can this DP be extended to multiple stocks?
  - Problem: Give the price info of 4 stocks over n days, can you find the best timings for “buy” and “sell” for each stock, such that the overall return is maximized?
- Other extensions
  - Different transaction fee rates for “buy” and “sell”?
  - Cash is better than stock?
    - 例如：一年至少保留30天擁有現金
  - Other reasonable constraints, please let me know!
    - Better make DP not applicable directly!

Our homework!

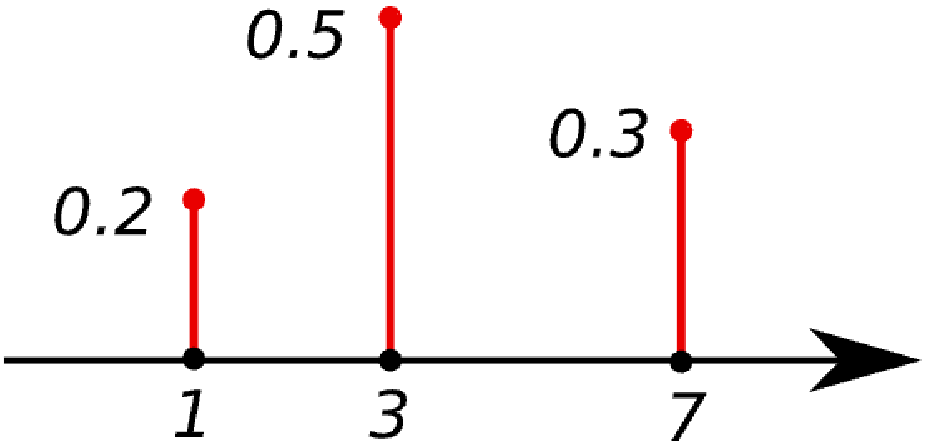
Basics about Random Variables

J.-S. Roger Jang (張智星)
MIR Lab, CSIE Dept.
National Taiwan University
jang@mirlab.org, http://mirlab.org/jang

2023/10/3

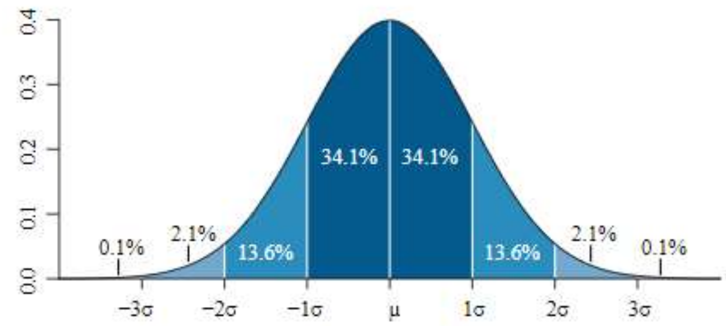
Math Background: Random Variables

- Discrete



Discrete Probability Disrtibution: This shows the probability mass function of a discrete probability distribution. The probabilities of the singletons {1}, {3}, and {7} are respectively 0.2, 0.5, 0.3. A set not containing any of these points has probability zero.

- Continuous



Probability Density Function: The image shows the probability density function (pdf) of the normal distribution, also called Gaussian or "bell curve", the most important continuous random distribution. As notated on the figure, the probabilities of intervals of values corresponds to the area under the curve.

https://courses.lumenlearning.com/boundless-statistics/chapter/discrete-random-variables/

Linear Combination of Random Variables

- Given two random variables X and Y
  - Definition
    - Mean: $\mu_X = E(X)$
    - Variance: $\sigma_X^2 = V(X) \triangleq E((X-\mu_X)^2) = E(X^2) - \mu_X^2$
    - Covariance: $\sigma_{XY} = \sigma_{YX} = E((X-\mu_X)(Y-\mu_Y)) = E(XY) - \mu_X\mu_Y$
  - Basic formulas for a single variable
    - $E(aX) = aE(X) \Rightarrow \mu_{aX} = a\mu_X$
    - $V(aX) = a^2V(X) \Rightarrow \sigma_{aX}^2 = a^2\sigma_X^2$
  - Extension to two variables (not necessarily independent)
    - $E(aX+bY) = aE(X) + bE(Y)$
      $\Rightarrow \mu_{aX+bY} = a\mu_X + b\mu_Y$
    - $V(aX+bY) = a^2V(X) + b^2V(Y) + 2abE((X-\mu_X)(Y-\mu_Y))$
      $\Rightarrow \sigma_{aX+bY}^2 = a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\sigma_{XY} = a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\sigma_X\sigma_Y\rho_{XY}$

https://hackmd.io/@rogerjang/BJwCkf0LK, with exercise!

Proof of Variance after Combination

- Proof of $\sigma^2_{aX+bY} = a^2\sigma^2_X + b^2\sigma^2_Y + 2ab\sigma_{XY} = a^2\sigma^2_X + b^2\sigma^2_Y + 2ab\sigma_X\sigma_Y\rho_{XY}$

$$
\begin{aligned}
V(aX+bY) &= E((aX+bY-E(aX+bY))^2) \\
&= E((aX+bY-a\mu_X-b\mu_Y))^2) \\
&= E((a(X-\mu_X)+b(Y-\mu_Y))^2) \\
&= E(a^2(X-\mu_X)^2+b^2(Y-\mu_Y)^2+2ab(X-\mu_X)(Y-\mu_Y)) \\
&= a^2E((X-\mu_X)^2)+b^2E((Y-\mu_Y)^2)+2abE((X-\mu_X)(Y-\mu_Y)) \\
&= a^2\sigma^2_X+b^2\sigma^2_Y+2ab\sigma_{XY} \\
&= a^2\sigma^2_X+b^2\sigma^2_Y+2ab\sigma_X\sigma_Y\rho_{XY}
\end{aligned}
$$

where $\rho_{XY} = \sigma_{XY}/(\sigma_X\sigma_Y)$ is the correlation coefficient between X and Y, with $-1 \le \rho_{XY} \le 1$.

https://hackmd.io/@rogerjang/BJwCkf0LK, with exercise!

Portfolio Optimization
投資組合最佳化

J.-S. Roger Jang (張智星)
MIR Lab, CSIE Dept.
National Taiwan University
jang@mirlab.org, http://mirlab.org/jang

2023/10/3

Outline

- Math background
  - Linear combination of random variables
- Portfolio optimization
  - Problem definition
  - Objective functions
  - Matrix formulas
  - Efficient frontier
- References

Linear Combination of Random Variables

- Given two random variables X and Y
  - Definition
    - Mean: $\mu_X = E(X)$
    - Variance: $\sigma_X^2 = V(X) \triangleq E((X-\mu_X)^2) = E(X^2) - \mu_X^2$
    - Covariance: $\sigma_{XY} = \sigma_{YX} = E((X-\mu_X)(Y-\mu_Y)) = E(XY) - \mu_X \mu_Y$
  - Basic formulas for a single variable
    - $E(aX) = aE(X) \Rightarrow \mu_{aX} = a\mu_X$
    - $V(aX) = a^2 V(X) \Rightarrow \sigma_{aX}^2 = a^2 \sigma_X^2$
  - Extension to two variables (not necessarily independent)
    - $E(aX+bY) = aE(X) + bE(Y)$
      $\Rightarrow \mu_{aX+bY} = a\mu_X + b\mu_Y$
    - $V(aX+bY) = a^2 V(X) + b^2 V(Y) + 2abE((X-\mu_X)(Y-\mu_Y))$
      $\Rightarrow \sigma_{aX+bY}^2 = a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\sigma_{xy} = a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\sigma_X\sigma_Y\rho_{XY}$

https://hackmd.io/@rogerjang/BJwCkf0LK, with exercise!

Portfolio Optimization (PO)

- Goal
  - To maximize the overall returns or minimize the overall variance of a portfolio of n assets based on each individual expected return and variance (aka risk or volatility)
- Facts
  - Introduced in a 1952 doctoral thesis by Harry Markowitz (awarded Nobel Memorial Prize in Economic Science in 1990)
  - Also known as mean-variance model or Markowitz model, which is foundational to Modern Portfolio Theory (MPT)
- Assumptions
  - Risk or volatility is equivalent to standard deviation.
  - No consideration for taxes, transaction fees, etc.

PO for Two Assets: Combined Mean and Variance

Given two risky assets as follows:

$$\begin{cases} \text{Asset 1: } \mu = \mu_1, \sigma^2 = \sigma_1^2 \\ \text{Asset 2: } \mu = \mu_2, \sigma^2 = \sigma_2^2 \end{cases}$$

We can use a weight vector $[w_1, w_2]^T$, with $w_1 + w_2 = 1$ to allocate these two assets to have mean return μ and risk (variance) σ^2:

μ: Overall mean return

σ: Overall risk or volatility

$$\begin{cases} \mu &= w_1\mu_1 + w_2\mu_2 \\ \sigma^2 &= w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\sigma_{12} \end{cases}$$

We can put the above equations into a matrix form (with $\sigma_{12} = \sigma_{21}$):

$$\begin{cases} \mu &= \begin{bmatrix} \mu_1 & \mu_2 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \\ \sigma^2 &= \begin{bmatrix} w_1 & w_2 \end{bmatrix} \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \end{cases}$$

PO for Two Assets: Efficient Frontier

Instead of using two parameters in the above expression, we can use only a single parameter w, with $w_1 = w$ and $w_2 = 1 - w$:

$$\begin{cases} \mu & = & w\mu_1 + (1-w)\mu_2 \\ \sigma^2 & = & w^2\sigma_1^2 + (1-w)^2\sigma_2^2 + 2w(1-w)\sigma_{12} \end{cases}$$

From the first expression, we have

$$w = \frac{\mu_2 - \mu}{\mu_2 - \mu_1}, 1 - w = \frac{\mu - \mu_1}{\mu_2 - \mu_1}$$

Therefore

$$\begin{aligned} \sigma^2 &= \left(\frac{\mu_2-\mu}{\mu_2-\mu_1}\right)^2\sigma_1^2 + \left(\frac{\mu-\mu_1}{\mu_2-\mu_1}\right)^2\sigma_2^2 + 2\left(\frac{\mu_2-\mu}{\mu_2-\mu_1}\right)\left(\frac{\mu-\mu_1}{\mu_2-\mu_1}\right)\sigma_{12} \\ &= \frac{1}{(\mu_2-\mu_1)^2}[(\mu^2 - 2\mu_2\mu + \mu_2^2)\sigma_1^2 + (\mu^2 - 2\mu_1\mu + \mu_1^2)\sigma_2^2 - 2(\mu^2 - (\mu_1+\mu_2)\mu + \mu_1\mu_2)\sigma_{12}] \\ &= \frac{1}{(\mu_2-\mu_1)^2}[(\sigma_1^2 + \sigma_2^2 - 2\sigma_{12})\mu^2 - 2(\mu_2\sigma_1^2 + \mu_1\sigma_2^2 - (\mu_1+\mu_2)\sigma_{12})\mu + \mu_2^2\sigma_1^2 + \mu_1^2\sigma_2^2 - 2\mu_1\mu_2\sigma_{12}] \end{aligned}$$

Since $\sigma_1^2 + \sigma_2^2 - 2\sigma_{12} \geq 2(\sigma_1\sigma_2 - \sigma_{12}) \geq 0$, the above equation is a hyperbola on the $\sigma - \mu$ plane. It can reduce to a parabola if $\sigma_1^2 + \sigma_2^2 - 2\sigma_{12} = 0$.

https://hackmd.io/@rogerjang/SJa16FaLt, with exercise!

PO for n Assets: Problem Definition

In general, for n assets, we can combine them to the overall return μ and risk σ:

$$\begin{cases} \mu & = & \boldsymbol{\mu}^T \mathbf{w} \\ \sigma^2 & = & \mathbf{w}^T \Sigma \mathbf{w} \end{cases}$$

where $\mathbf{w} = [w_1, \ldots, w_n]^T$, $\boldsymbol{\mu} = [\mu_1, \ldots, \mu_n]^T$, and Σ is the covariance matrix of these n assets.

Suppose we want to minimize risk with fixed return, as follows.

$$\min_{\mathbf{w}} \sigma^2 = \mathbf{w}^T \Sigma \mathbf{w}$$

$$s.t. \begin{cases} \mathbf{1}^T \mathbf{w} = 1 \\ \boldsymbol{\mu}^T \mathbf{w} = \mu_0 \end{cases}$$

where $\mathbf{1} = [1, \ldots, 1]^T$.

Objective Functions for PO (1/2)

- Minimize risk with fixed return: Given a return μ, find the weights to minimize the overall variance σ^2. (給定預期報酬值，最佳投資組合將產生最小風險。)

$$\min_{\mathbf{w}} \sigma^2 = \mathbf{w}^T \Sigma \mathbf{w}$$

$$s.t. \begin{cases} \mathbf{w}^T \mathbf{1} = 1 \\ \boldsymbol{\mu}^T \mathbf{w} = \mu_0 \end{cases}$$

where $\mathbf{1} = [1, \ldots, 1]^T$.

- Maxmimize return with fixed risk: Given a variance σ^2, find the weights to maximize the overall return μ. (給定風險下，最佳投資組合將產生預期報酬最大值。)

$$\max_{\mathbf{w}} \mu = \boldsymbol{\mu}^T \mathbf{w}$$

$$s.t. \begin{cases} \mathbf{w}^T \mathbf{1} = 1 \\ \mathbf{w}^T \Sigma \mathbf{w} = \sigma_0^2 \end{cases}$$

Objective Functions for PO (2/2)

- Minimize risk regardless of return: Find the weights to minimize the overall variance σ^2 regardless of the return. (讓最佳投資組合將產生最小風險，而完全不看報酬。)

$$\min_{\mathbf{w}} \sigma^2 = \mathbf{w}^T \Sigma \mathbf{w}$$
$$s.t. \ \mathbf{w}^T \mathbf{1} = 1.$$

- Maximize the Sharpe ratio

$$\max_{\mathbf{w}} \frac{\mu - \mu_0}{\sigma}$$
$$s.t. \ \mathbf{w}^T \mathbf{1} = 1$$

- Maximize the difference between returm and risk

$$\max_{\mathbf{w}} \mu - \beta\sigma$$
$$s.t. \ \mathbf{w}^T \mathbf{1} = 1$$

In fact, there are a lot more objective functions and constraints in practice!

$PO_{n=2}$: ρ_{12}=1

$$\begin{cases} \text{Asset 1: } \mu = \mu_1, \sigma = \sigma_1 \\ \text{Asset 2: } \mu = \mu_2, \sigma = \sigma_2 \end{cases} \xrightarrow{w_1 + w_2 = 1} \begin{cases} \mu & = & w_1\mu_1 + w_2\mu_2 \\ \sigma^2 & = & w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\sigma_1\sigma_2\rho_{12} \end{cases}$$

When $\rho_{12} = 1$, we have

$$\sigma^2 = w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\sigma_1\sigma_2 = (w_1\sigma_1 + w_2\sigma_2)^2 \Rightarrow \begin{cases} \mu & = & w_1\mu_1 + w_2\mu_2 \\ \sigma & = & |w_1\sigma_1 + w_2\sigma_2| \end{cases}$$

As w_1 is changing from 0 to 1, the above equations represent a line connecting (σ_1, μ_1) (when $w_1 = 1$ and $w_2 = 0$) and (σ_2, μ_2) (when $w_1 = 0$ and $w_2 = 1$). So the minimum variance is $\min(\sigma_1^2, \sigma_2^2)$.

$PO_{n=2}$: ρ_{12}=0

When $\rho_{12} = 0$, we have

$$\sigma^2 = w_1^2\sigma_1^2 + w_2^2\sigma_2^2$$

By using Cauchy-Schwartz inequality, we have

$$(w_1^2\sigma_1^2 + w_2^2\sigma_2^2)(\sigma_1^{-2} + \sigma_2^{-2}) \geq (w_1 + w_2)^2 = 1$$

Therefore the minimum variance can be derived as follows:

$$\sigma^2 = w_1^2\sigma_1^2 + w_2^2\sigma_2^2 \geq (\sigma_1^{-2} + \sigma_2^{-2})^{-1} = \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2 + \sigma_2^2}$$

The equality holds when

$$w_1^2\sigma_1^2/\sigma_1^{-2} = w_2^2\sigma_2^2/\sigma_2^{-2} \Rightarrow w_1 = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}, w_2 = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}$$

$PO_{n=2}$: ρ_{12}=-1

When $\rho_{12} = -1$, we have

$$\sigma^2 = w_1^2\sigma_1^2 + w_2^2\sigma_2^2 - 2w_1w_2\sigma_1\sigma_2 = (w_1\sigma_1 - w_2\sigma_2)^2$$

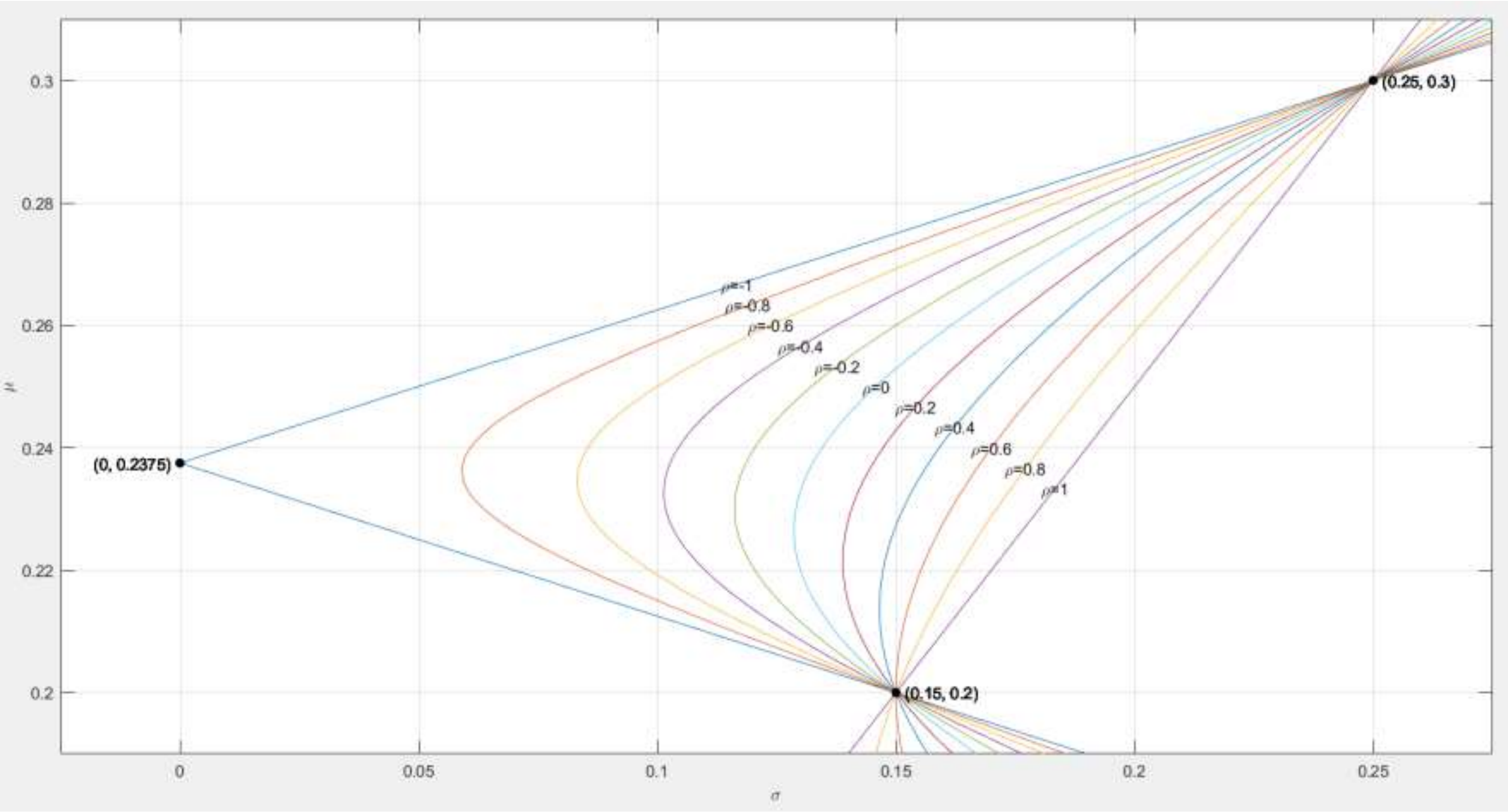
In this case, we can achieve zero risk by setting

$$w_1\sigma_1 = w_2\sigma_2 \Rightarrow w_1 = \frac{\sigma_2}{\sigma_1 + \sigma_2}, w_2 = \frac{\sigma_1}{\sigma_1 + \sigma_2}$$

Therefore the minimum variace is 0, and the corresponding return is $\frac{\sigma_2\mu_1+\sigma_1\mu_2}{\sigma_1+\sigma_2}$.

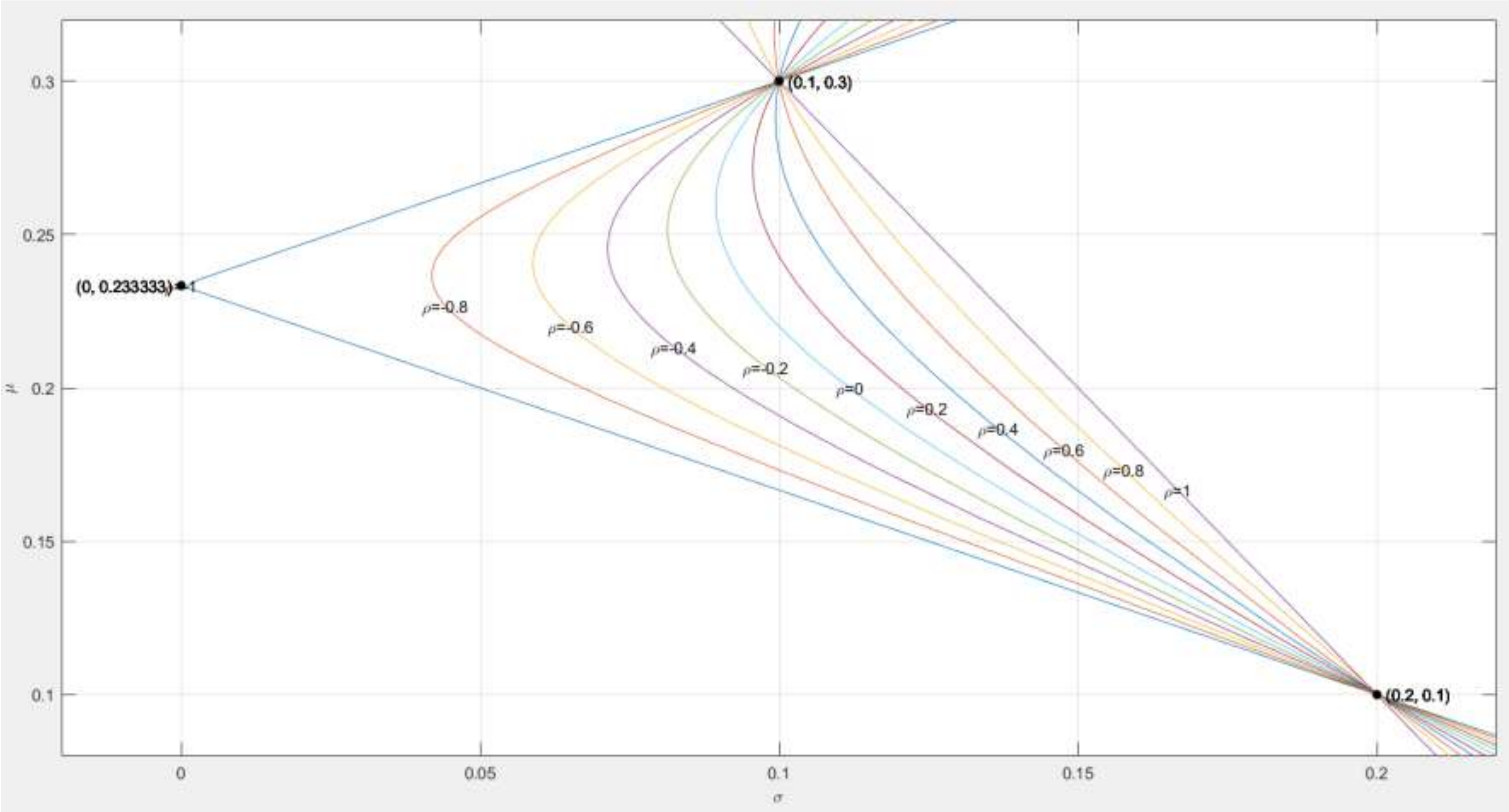
$PO_{n=2}$: Efficient Frontier with Varying ρ_{12} (1/2)

- n=2: (σ_1, μ_1)=(0.15, 0.2) and (σ_2, μ_2)=(0.25, 0.3)



$PO_{n=2}$: Efficient Frontier with Varying ρ_{12} (2/2)

- n=2: (σ_1, μ_1)=(0.2, 0.1) and (σ_2, μ_2)=(0.1, 0.3)



PO_n: Min. Variance Only

In general, for n assets, we can combine them to the overall return μ and risk σ:

$$\begin{cases} \mu & = & \boldsymbol{\mu}^T \mathbf{w} \\ \sigma^2 & = & \mathbf{w}^T \Sigma \mathbf{w} \end{cases}$$

where $\mathbf{w} = [w_1, \ldots, w_n]^T$, $\boldsymbol{\mu} = [\mu_1, \ldots, \mu_n]^T$, and Σ is the covariance matrix of these n assets.

Suppose we want to minimize the overall risk regardless of the overall return, then the problem can be formulated as follows:

$$\min_{\mathbf{w}} \sigma^2 = \mathbf{w}^T \Sigma \mathbf{w}$$
$$\text{s.t. } \mathbf{1}^T \mathbf{w} = 1$$

PO_n: Min. Variance Only (Block-form Solution)

To find the solution to this constrained optimization problem, we can formulate a new objective function using the Lagrange multiplier:

$$\max_{\mathbf{w},\lambda} J(\mathbf{w},\lambda) = \mathbf{w}^T \Sigma \mathbf{w} + \lambda(\mathbf{1}^T \mathbf{w} - 1).$$

By taking the gradient and set it to zero, we have

$$\nabla_{\boldsymbol{w}} J(\boldsymbol{w},\lambda) = 2\Sigma\mathbf{w} + \mathbf{1}\lambda = 0 \Rightarrow \mathbf{w} = -\frac{1}{2}\Sigma^{-1}\mathbf{1}\lambda$$

Since $\mathbf{1}^T\mathbf{w} = 1$, we have

$$-\frac{1}{2}\mathbf{1}^T\Sigma^{-1}\mathbf{1}\lambda = 1 \Rightarrow \lambda = -\frac{2}{\mathbf{1}^T\Sigma^{-1}\mathbf{1}}.$$

Therefore

$$\mathbf{w} = \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}^T\Sigma^{-1}\mathbf{1}}$$

PO_n: Min. Var. with Fixed Return

To solve this problem, we can use the Lagrange multiplier to form a new objective function:

$$\max_{\mathbf{w},\lambda_1,\lambda_2} J(\mathbf{w},\lambda_1,\lambda_2) = \mathbf{w}^T\Sigma\mathbf{w} + \lambda_1(\mathbf{1}^T\mathbf{w}-1) + \lambda_2(\boldsymbol{\mu}^T\mathbf{w}-\mu_0),$$

We can take the gradient and set it to zero to have the following equations:

$$\begin{cases} 2\Sigma\mathbf{w} + \mathbf{1}\lambda_1 + \boldsymbol{\mu}\lambda_2 &= \mathbf{0} \\ \mathbf{1}^T\mathbf{w} &= 1 \\ \boldsymbol{\mu}^T\mathbf{w} &= \mu_0 \end{cases}$$

(Note that we omit the use of "hat" to keep simplicity.)

When n=3

$$\begin{bmatrix} 2\Sigma_{11} & 2\Sigma_{12} & 2\Sigma_{13} & 1 & \mu_1 \\ 2\Sigma_{21} & 2\Sigma_{22} & 2\Sigma_{23} & 1 & \mu_2 \\ 2\Sigma_{31} & 2\Sigma_{32} & 2\Sigma_{33} & 1 & \mu_3 \\ 1 & 1 & 1 & 0 & 0 \\ \mu_1 & \mu_2 & \mu_3 & 0 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ \mu_0 \end{bmatrix}$$

PO_n: Min. Var. with Fixed Return (Block-form Solution)

If we put the above equations into the block form:

$$\begin{bmatrix} 2\Sigma & B \\ B^T & \mathbf{0}_{2\times 2} \end{bmatrix} \begin{bmatrix} \mathbf{w} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{n\times 2} \\ \mathbf{k} \end{bmatrix}$$

where $B = \begin{bmatrix} 1 & \mu_1 \\ 1 & \mu_2 \\ \vdots & \vdots \\ 1 & \mu_n \end{bmatrix}$, $\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$, $\boldsymbol{\lambda} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}$, and $\mathbf{k} = \begin{bmatrix} 1 \\ \mu_0 \end{bmatrix}$.

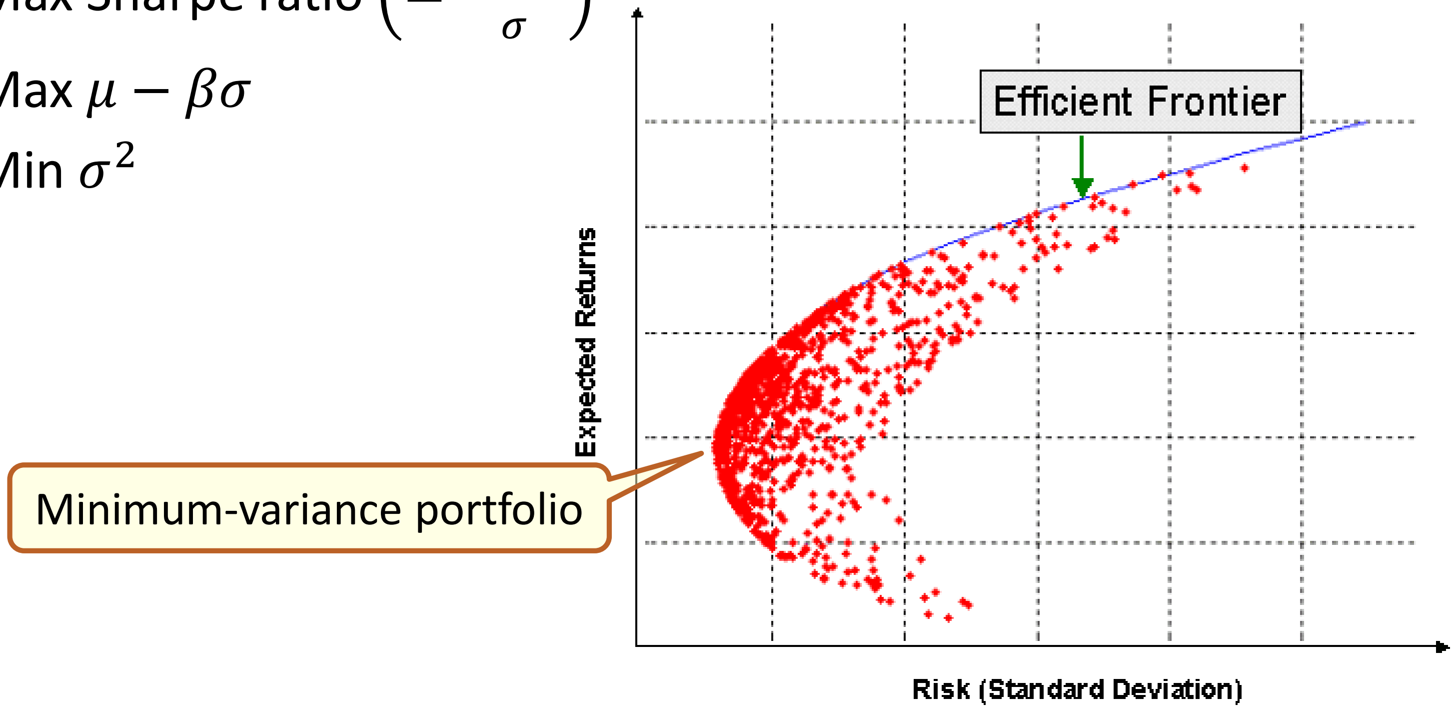
By direct matrix manipulation, we can obtain the solution as follows:

$$\begin{cases} \mathbf{w} & = & \Sigma^{-1}B(B^T\Sigma^{-1}B)^{-1}\mathbf{k} \\ \lambda & = & -2(B^T\Sigma^{-1}B)^{-1}\mathbf{k} \end{cases}$$

https://hackmd.io/@rogerjang/BkXxCdCIF, with exercises!

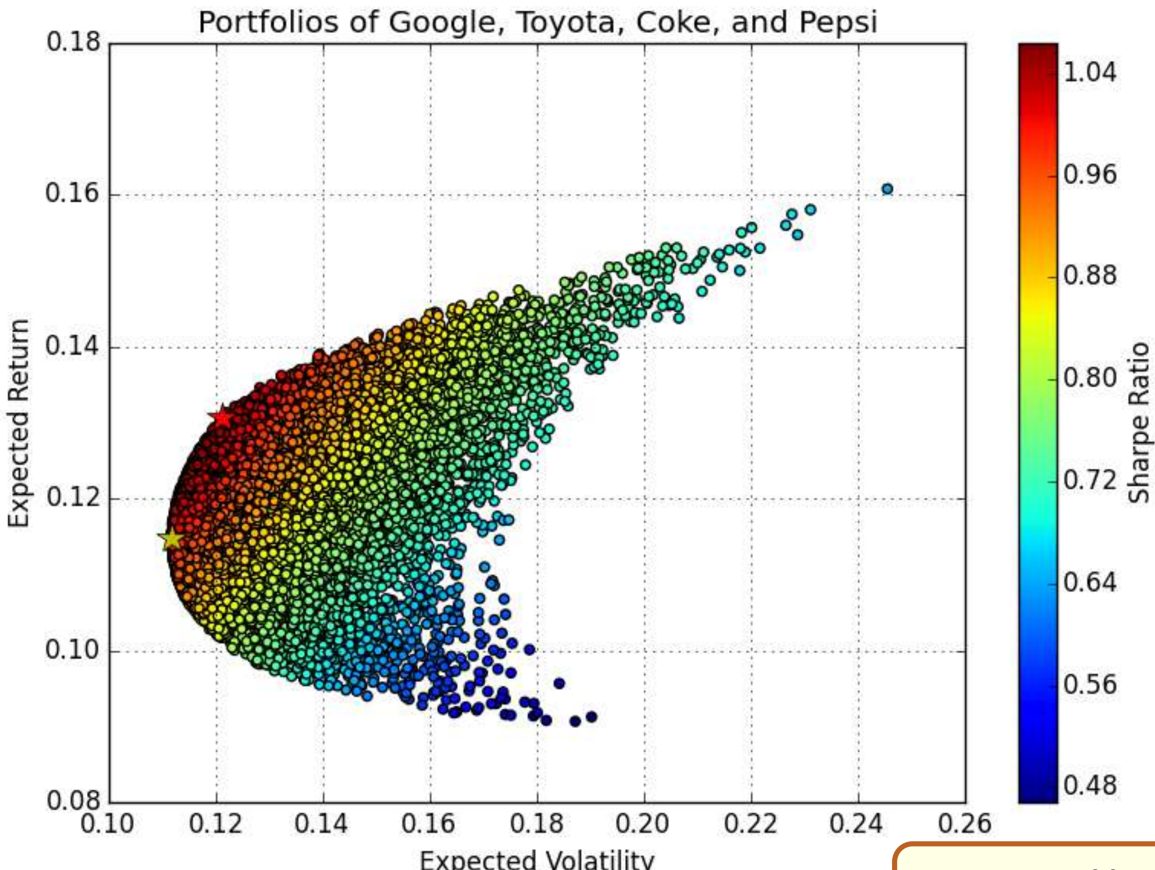
Efficient Frontier

- Efficient frontier for portfolio optimization
  - Max μ (return) with fixed σ^2 (risk)
  - Min σ^2 (risk) with fixed μ (return)
  - Max Sharpe ratio $\left(=\frac{\mu-\mu_f}{\sigma}\right)$
  - Max $\mu - \beta\sigma$
  - Min σ^2



Resources

- Investment Portfolio Optimization (with Python code)



More references:
https://hackmd.io/@rogerjang/SJN4FQbvF

Other Things to Consider

- How to compute $\boldsymbol{\mu}$ (returns) and $\sum$ (covariance matrix)?
- When to rebalance the assets?
- Other constraints
  - Max. value of n
  - Max. number of changes in assets
  - Conversion of individual risk attributes to objective functions
  - Conversion of individual preferences to objection functions

Exercises (1/2)

1. In PO of $n = 2$, when will the efficient frontier reduce to a straight line?
2. In PO of $n = 2$, when will the efficient frontier reduce to a parabola?
3. In PO of $n = 2$, when will the overall risk go to zero? What are the weights when this happens?
4. In PO of $n = 2$, can you derive the general formula for minimum-variance portfolio?
   - What is the minimum variance?
   - What is the corresponding return and weights?
5. Given two risky assets as follows:

$$\begin{cases} \text{Asset 1: } \mu = 0.2, \sigma = 0.1 \\ \text{Asset 2: } \mu = 0.3, \sigma = 0.2 \end{cases}$$

   Under the following conditions, what are the corresponding minimum variances when we achieve minimum-variance portfolio?
   - $\rho_{12} = 1$
   - $\rho_{12} = 0$
   - $\rho_{12} = -1$

Exercises (2/2)

6. Given two risky assets as follows:

$$\begin{cases} \text{Asset 1: } \mu = 0.2, \sigma = 0.1 \\ \text{Asset 2: } \mu = 0.3, \sigma = 0.2 \end{cases}$$

And the correlation coefficient of these two assets is $\rho_{12} = 0$. We want to perform portfolio optimization with investment weigthing of w_1 and w_2 for assets 1 and 2, respectively.

- What are the overall μ (return) and σ (volatility) when $w_1 = 0.4$ and $w_2 = 0.6$?
- What are the overall μ, overall σ, and w_1 for achieving the minimum-variance portfolio?

7. Given two risky assets as follows:

$$\begin{cases} \text{Asset 1: } \mu = 0.2, \sigma = 0.1 \\ \text{Asset 2: } \mu = 0.3, \sigma = 0.2 \end{cases}$$

And the correlation coefficient of these two assets is $\rho_{12} = 0.4$. We want to perform portfolio optimization with investment weigthing of w_1 and w_2 for assets 1 and 2, respectively.

- What are the overall μ (return) and σ (volatility) when $w_1 = 0.4$ and $w_2 = 0.6$?
- What are the overall μ, overall σ, and w_1 for achieving the minimum-variance portfolio?

References

- References at hackmd (with detailed math formula)
  - Intro to portfolio optimization
  - Objective functions for portfolio optim.
  - Portfolio for 2 assets
  - Portfolio optim.: Min. risk only
  - Portfolio optim.: Min. risk with fixed return
  - References