Investment: A Brief Note

Shuhua Xiao

更新：*2022 Fall; Note for **Investment** Course (Master).*

目录

0 引论

0.1 投资与投资学

0.1.1 投资

投资是为了获得可能但并不确定的未来值而作出牺牲确定的现值的行为（William F. Sharpe，1990 年获得诺贝尔经济学奖）。

- 时间性：牺牲当前消费；增加未来消费；资金的时间价值
  关于时间价值：CAPM: $E(r_i) - r_f = \beta_i(E(r_M - r_f))$，$r_f$ 是时间价值最基本的代表
  指数模型：$r_i - r_f = \alpha_i + \beta_i(r_m - r_f) + \epsilon_i$
- 收益性：增加投资者的财富来满足未来的消费
- 风险性：损益的不确定性

0.1.2 投资学

研究投资行为及均衡定价的科学。投资学是金融学的核心课程之一。

0.1.3 投资学课程的基本框架

三大内容：金融工具与金融市场、投资理论、证券投资分析实务。

- 金融工具与金融市场：基础性金融工具与衍生工具、交易所、中间商、以及市场微观结构等。（金融市场学）
- 投资理论：证券的风险与收益、组合投资理论、CAPM 理论、APT、有效市场理论、债券理论、期权定价模型等。（投资学）
- 投资实务：基本面分析和技术分析。（证券投资学）

0.2 金融资产

0.2.1 财富与资产

- 财富（Wealth）：所拥有的产品与服务以及可以获得产品与服务的等价物。
  房子、车子、面包、现金、股票等。
- 资产（Assets）：所有可以创造财富的东西。

厂房、土地、专利技术、银行存款、股票、债券等。

0.2.2 实物资产与金融资产

- 实物资产（Real assets）：直接创造财富的资产。包括土地、建筑、机器、知识等。代表一个经济的生产能力，决定一个社会的财富。
- 金融资产（Financial assets）：实物资产的要求权，定义实物资产在投资者之间的配置。
  金融资产的价值与其物质形态没有任何关系：股票本身可能并不比印制股票的纸张更值钱。
  整个社会财富的总量与金融资产数量无关，金融资产不是社会财富的代表。

0.2.3 金融资产在经济中的作用

- 消费的时机安排: 个人现实消费与现实收入分离，将高收入期的购买力转移到低收入期。
- 风险的分配: 风险的分散、分担和优化配置。
- 所有权与经营权分离

0.3 金融市场

0.3.1 金融市场

：金融市场是金融资产的融通场所。

四大分类：

- 合约性质：债券市场、股票市场、期货市场、期权市场。
- 期限长短：货币市场和资本市场
- 功能：初级（一级）市场——发行市场，二级市场——交易市场。
- 组织结构：交易所、场外市场等

0.3.2 金融市场的主体

（1）家庭部门：既是金融市场资金的主要供给者，又是投资者。

需求多样性：

寻求高收益工具

寻求保值

高负税的投资者寻求免（低）税的金融工具

风险对冲

(2) 企业：融资

间接融资：向银行借款

直接融资：直接向家庭借款（发行股票、债券等）

目标：以低成本卖出高价格（一级市场）的证券。

从理论上说，企业可以直接从家庭借款，但实际上需要通过专门的金融机构设计融资方案——投资经济分析向企业提出建议，是发行股票还是债券、利息多少、期限多少等。

(3) 政府

弥补财政赤字

实施货币政策：政府可以通过公开市场业务控制短期国债的投放量来控制利率。

规范金融环境：政府是监管者、同时管制又是金融创新的动力。

0.3.3 金融机构

(1) 金融中介：为间接融资提供服务

商业银行、保险公司、投资公司、共同基金、信托机构等。

(2) 证券业机构：为直接融资提供服务

投资银行；交易所

0.3.4 金融市场的主要功能

(1) 投资、融资：投融资是一个问题的两面。

(2) 增强资产的流动性：实物资产的证券化使其具有可分割性、易转让性。

(3) 优化资源配置，提升经济运行效率：使资源向有效率的企业转移。

1 最优化基础

1.1 Optimization

1.1.1 Optimization model

Choose the best solution under some constraints

如果 $f(x)$ 是一个凸函数，$g_i(x)$ 是一个凸函数，$h_j(x)$ 是一个线性函数，那么这就是一个凸优化问题。

$$\begin{aligned} &\min_x f(x) \\ &\text{s.t. } g_i(x) \le 0, i = 1, \ldots, p \\ &\quad h_j(x) = 0, j = 1, \ldots, q \end{aligned} \tag{P}$$

决策变量：$x \in R^n$

目标函数：$f(x)$

约束：$g_i(x) \le 0, h_j(x) = 0$。If $p = q = 0$, then (P) is called an unconstrained optimization problem（无约束优化问题）.

$\min -2x_1 - 3x_2$ $x_1 + 2x_2 - 8 \le 0$ $4x_1 - 16 \le 0$ $4x_2 - 12 \le 0$ $-x_1, -x_2 \le 0$

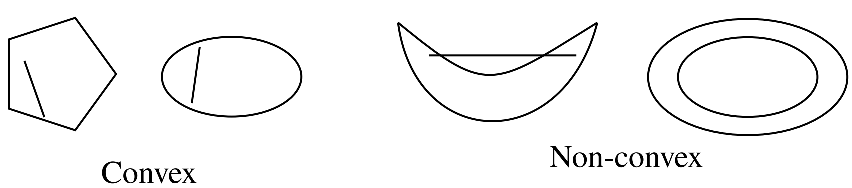
在凸集上，凸函数取极小、凹函数取极大，都是凸优化问题。

1.1.2 Convexity 凸性

Convex set

Suppose that $S \in R^n$ and for any given（任意给定）$x^1, x^2 \in S$ and $\lambda \in [0, 1]$,we must have

$$\lambda x^1 + (1 - \lambda)x^2 \in S$$



若不限制 λ，则满足前述定义的组合为仿射组合。

1.1.3 Optimality conditions 最优性条件

凸函数在凸集上求极小，是为凸优化问题。**Gradient 梯度**

多元函数的一阶求导信息

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \frac{\partial f(x)}{\partial x_2} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{bmatrix} \tag{1}$$

Hessian matrix 海森矩阵

多元函数的二阶求导信息

$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix} \tag{2}$$

梯度的方向是上升方向，负梯度的方向是下降方向。

Taylor Expansion

Univarite（一维情况下）:

Δx 表示从当前的 x^0 点往前探索一点地方。

$$f\left(x^0 + \Delta x\right) = f\left(x^0\right) + f'\left(x^0\right)\Delta x + \frac{1}{2} f''\left(x^0\right)\Delta x^2 + o\left(\Delta x^2\right)$$

Multivariate（多维情况下）:

$$f\left(x^0 + \Delta x\right) = f\left(x^0\right) + \nabla f\left(x^0\right)^T \Delta x + \frac{1}{2}\Delta x^T \nabla^2 f\left(x^0\right)\Delta x + o\left(\Delta x^T \Delta x\right)$$

下降方向 Descent direction

x^* 的下降方向：$d \in R^n$ such that $\nabla f(x^*)^T d < 0$

Steepset descent direction 最速下降法:

因为:

$$f(x^0 + \Delta x) - f(x^0) \approx \nabla f(x^*)^T d < 0$$

所以 $f(x^0 + \Delta x)$ 更小，要往 Δx 方向继续走，所以我们说下降方向。可以在下降方向前加一个参数，视为步长。

$$d = -\nabla f(x^*) \xrightarrow{\nabla f(x^*)^T \neq 0} -\nabla f(x^*)^T \nabla f(x^*) < 0$$

为什么有箭头的变换？因为乘以个正定矩阵，依旧是下降方向。取特殊一点取了雅可比矩阵而已。牛顿下降法就是把你乘的这个矩阵搞得更特殊一点，换成 hessian 矩阵。

负梯度

Newton descent direction 牛顿法（牛顿下降方向）：

$$d = -\left[\nabla^2 f\left(x^0\right)\right]^{-1} \nabla f(x^*) \xrightarrow[\left[\nabla^2 f\left(x^0\right)\right]^{-1} \succ 0]{\nabla f(x^*)^T \neq 0} -\nabla f(x^*)^T \left[\nabla^2 f\left(x^0\right)\right]^{-1} \nabla f(x^*) < 0$$

海森矩阵(正定的)的逆乘上负梯度。因为海森矩阵是正定的，所以 $\nabla f(x^*)^T \left[\nabla^2 f\left(x^0\right)\right]^{-1} \nabla f(x^*)$ 是一个正定二次型。

如果有约束，把梯度下降的方向往边界投影一下就可以做到了。(跟单纯形法有类似的意思，就把梯度下降的方向往边界投影一下就可以往那边儿走)

在这个一阶方法和二阶方法之间有一个拟牛顿法。拟牛顿法只需一阶信息。具体地，是计算两个点的梯度，做差，来近似拟合 Hessian 矩阵。

First-order optimality condition, FOC, Unconstrained optimization problem :

If x^* is a local optimizer to (P), then

$$\nabla f(x^*) = 0$$

If (P) is convex, then it is also a sufficient condition for the global optimality.

对于一个无约束优化问题，x^* 是优化问题 (P) 的局部最优解的话，那么一定满足 $\nabla f(x^*) = 0$。如果不为 0，它一定可以找到一个下降方向（负梯度方向、牛顿方向都可以）；如果 (P) 是凸函数，那么上述是充分必要条件。

而对于约束优化问题

需要考虑有效约束（blinding）

Denote

$$I(x^*) = \{i | g_i(x^*) = 0, i \in 1, 2, \cdots, p\}$$

记使得 $g_i(x^*) = 0$ 的点的集合为 $I(x^*)$，其他的不管，因其他的不是有效的约束

Feasible direction 可行方向 of x^* : $d \in R^n$ such that

$$\nabla g_i(x^*)^T d \le 0, i \in I(x^*)$$

and

$$\nabla h_j(x^*)^T d = 0, j = 1, 2, \cdots, q.$$

上述二式是下述式的变形：

$$g(x^0 + \Delta x) - g(x^0) \approx \nabla g(x^*)^T d \le 0$$

考虑到 $g(x^0) = 0$，则有：

$$g(x^0 + \Delta x) \approx \nabla g(x^*)^T d \le 0$$

$$h(x^0 + \Delta x) - h(x^0) \approx \nabla h(x^*)^T d = 0$$

考虑到 $h(x^0) = 0$，则有：

$$h(x^0 + \Delta x) \approx \nabla h(x^*)^T d \le 0$$

结合下降方向和可行方向，二者若有交集，则为可行下降方向。

$$\left\{ \begin{array}{cc} & \nabla f\left(x^*\right)^T d < 0 \\ d: & \nabla g_i\left(x^*\right)^T d \le 0, i \in I\left(x^*\right) \\ & \nabla h_j\left(x^*\right)^T d = 0, j = 1, 2, \ldots, q \end{array} \right\} = \phi \rightarrow optimality\ condition$$

可行方向：在这个方向上走，不会走出可行域。ϕ 是可行下降方向的集合。这个集合是由目标函数的梯度和约束的梯度刻画的。如果 ϕ 是空集，意味着已经达到了（局部）最优解。

ϕ 是空集，再加上 Slater condition，等价于下述的 KKT 条件。（如果不加 Slater condition，则 f(x) 的梯度前面有一个参数 λ 且其可以为 0）

Kuhn-Tucker condition: If x^* is a local optimizer to (P), then there exist $\lambda_i \ge 0, i \in I(x^*)$ and $\mu_j, j = 1, 2, \cdots, q$ such that

$$\nabla f\left(x^*\right) + \sum_{i \in I(x^*)} \lambda_i \nabla g_i\left(x^*\right) + \sum_{j=1}^{q} \mu_j \nabla h_j\left(x^*\right) = 0$$

If (P) is convex, then it is also a sufficient condition for the global optimality. (need some regular condition, e.g. Slater condition) 如果 (P) 是凸的，那么它也是全局最优性的充分条件。

(是必要条件时，需要一些常规条件，例如 Slater 条件、紧约束的梯度是线性无关的；是充分条件时，则不需要)

前述 $i \in I(x^*)$ 是意味着只取/只考虑有效的（紧的）不等式约束。

Slater condition: the feasible set is convex and $\exists \tilde{x}, g_i(\tilde{x}) < 0, i = 1, \ldots, p$.

等价条件 equivalent formulation:

$$\nabla f\left(x^*\right)+\sum_{i=1}^p \lambda_i \nabla g_i\left(x^*\right)+\sum_{j=1}^q \mu_j \nabla h_j\left(x^*\right)=0$$

$$\lambda_i g_i\left(x^*\right)=0, \lambda_i \geq 0, i=1,2, \ldots, p$$

目标函数的梯度可以表示为约束函数的梯度的线性组合。

$\lambda_i g_i\left(x^*\right)=0, \lambda_i \geq 0, i=1,2, \ldots, p$ 是互补松弛条件。相当于不紧的约束（$g_i(x) > 0$）前面的 λ_i 等于 0。

1.1.4 Duality 对偶

Lagrange functon:

$$L(x,\lambda,\mu) = f(x) + \sum_{i=1}^{n} \lambda_i g_i(x) + \sum_{j=1}^{q} \mu_j h_j(x)$$

λ, μ: Dual variables / Lagrange multipliers

Dual function:

$$g(\lambda,\mu) = \min_x L(x,\lambda,\mu) \quad (\lambda \geq 0)$$

上述为拉格朗日对偶函数。对这个函数再求极大，就是对偶问题，这是一个凸优化问题。这儿需要再求一个笔记

Convexity 对偶问题的凸性: Dual function $g(\lambda,\mu)$ is a concave function even when (P) is not convex.

Can be proved by definition easily (if $-f(x)$ is convex then $f(x)$ is concave)

Lower bound: Dual function yields a lower bound to the optimal value p^* of (P). 对偶问题可以得到原问题最优解的下界。

For any $\lambda \geq 0$, μ and **any** feasible solution $\tilde{x}$ to (P)

$$\sum_{i=1}^{n} \lambda_i g_i(\tilde{x}) + \sum_{j=1}^{q} \mu_j h_j(\tilde{x}) \leq 0$$

$$L(\tilde{x},\lambda,\mu) = f(\tilde{x}) + \sum_{i=1}^{n} \lambda_i g_i(\tilde{x}) + \sum_{j=1}^{q} \mu_j h_j(\tilde{x}) \leq f(\tilde{x})$$

$$g(\lambda,\mu) = \min_x L(x,\lambda,\mu) \leq L(\tilde{x},\lambda,\mu) \leq f(\tilde{x})$$

第一个式子是小于零的项加等于 0 的项，结果小于 0；Inequality holds for any feasible solution $\tilde{x}$, thus $g(\lambda,\mu) \leq p^*$.

Dual Problem:

$$(DP) \quad \max_{\lambda \geq 0, \mu} g(\lambda,\mu) = \max_{\lambda \geq 0, \mu} \min_x L(x,\lambda,\mu)$$

- **Weak Duality**:Denote p^* as the optimal value to (P) and d^* the optimal value to (DP). Then $d^* \leq p^*$.

- **Strong Duality**: If (P) is a convex program 凸. Under some regular condition (e.g. Slater condition), we have $d^* = p^*$. 零对偶间隙。
  线性规划的原问题和对偶问题都是凸问题。不需要 Slater condition 也是强对偶。

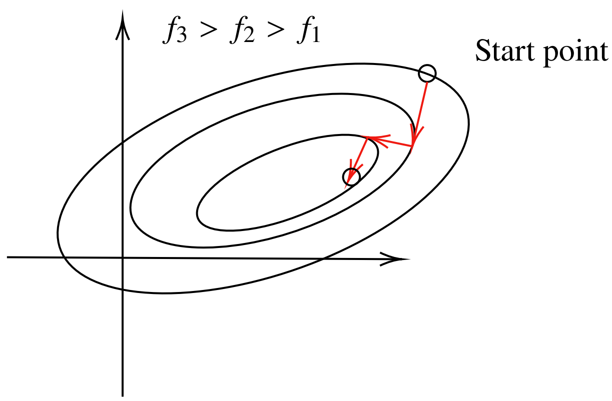
1.1.5 Algorithm 算法

Basic idea: Search along the **feasible descent direction** 沿着可行下降方向去找

Steepest descent method(unconstrained problem// 举例：无约束优化问题的最速下降法)

- Step1: Choose initial point x^0 and tolerance $\varepsilon > 0$, set $k := 0$
- Step2: Compute $\nabla f(x^k), \left\|\nabla f(x^k)\right\| \le \varepsilon$, stop, output x^k, else go to step 3
- Step3: Set $d^k = -\nabla f(x^k)$ 负梯度方向作为下降方向
- Step4: Compute t^k, such that $f(x^k + t_k d^k) = \min_{t\ge 0} f(x^k + td^k)$, set , go to step 2

Other methods: Newton,Quasi-Newton 等



如果是 Apply the descent method to a sequence of parameterized penalty functions (unconstrained optimization). 约束问题，可以把约束放到目标函数里作为一个惩罚项，不满足的话罚值就很大 → 罚函数方法（见下）

Penalty function method(constrained problem)

$$\min_x F_{c_k}(x) = f(x) + p_{c_k}(x)$$

with $p_{c_k}(x) = c_k \sum_{i=1}^{p} [\max(g_i(x), 0)]^2 + \frac{c_k}{2} \sum_{j=1}^{q} \left[h_j(x)\right]^2$

具体操作里，c_k 是逐渐变大的，$c_k \to +\infty \Rightarrow x_k^* \to x^*$

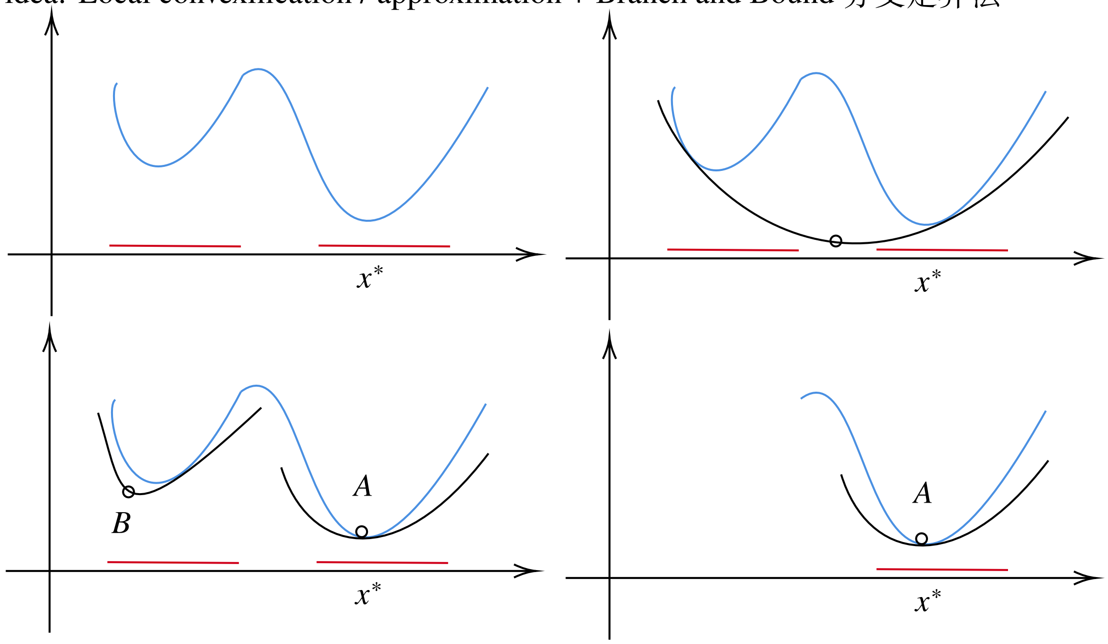
idea: Apply the descent method to a sequence of parameterized penalty functions (unconstrained optimization).

Other methods for constrained problems:

- Sequential approximations: SQP, Frank-Wolfe, gradient projection,... 每个梯度的时候都用线性逼近 不太好 序列二次规划 (二次逼近)
- Duality based methods: primal-dual iteration, alternative direction method(交替方向法), . . .
- Interior point methods 内点法 (convex program): Newton’s method applied to K-T system, ···对于凸优化问题，内点法是很好的方法 (就是用牛顿法去解 KT 条件)

Global optimization: non-convex problem 对于非凸优化问题，怎么求解呢?

idea: Local convexification / approximation + Branch and Bound 分支定界法



对上面这个非凸函数做一个放松 + 集合也搞大一点 那就可以找到最优解如果这个问题是原问题的可行解，那这个问题就结束了已经；如果不是的话，那它也为原问题提供了一个下界；没有可行解，那我就分块一下，作凸优化；A 比 B 好，B 那边就没有希望了，不用考虑（分支）。

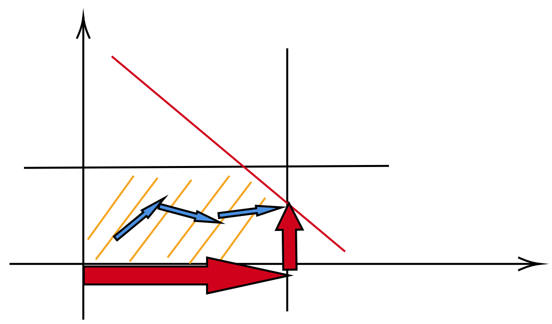
1.1.6 Conic optimization 锥优化

SDP 本身也是一个凸优化问题

(1)LP（Linear Programming）

$$
\begin{aligned}
&\min c^T x \\
&\text{s.t. } Ax - b \geq 0
\end{aligned}
$$

对于 LP（Linear Programming）问题，可以用单纯形法（Simplex method, Dantzig 1947）和内点法（Interior point method, Karmarkar 1984）进行计算。单纯形法可以计算后，下一次从那个点继续开始计算；内点法要从头开始。



(2)Second-Order Cone Programming(SOCP) 二阶锥规划

我们可以写出线性规划的一般形式，也就是:

$$
\begin{aligned}
&\min c^T x \\
&\text{s.t. } Ax - b \geq 0
\end{aligned}
$$

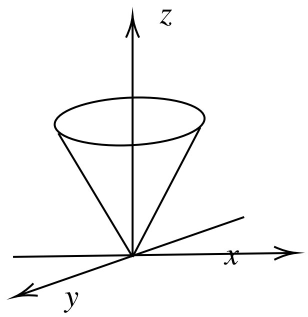
其中，$Ax - b$ 是一个向量，可以用 z_1，z_2 等来代替。换言之，$Ax - b > 0$ 也就是向量的每个分量大于等于 0。那么我们就有 $z \geq 0$，也即 $z_i \geq 0, i = 1, \cdots, n$。

如果类比到 SOCP，我们可以写出:

$$
\begin{aligned}
&\min c^T x \\
&\text{s.t. } Ax - b \geq_{socp} 0
\end{aligned}
$$

其中，$z = (z_1, z_2, \cdots, z_n)^T$。类比一下，也就有 $z \geq_{socp} 0 \Leftrightarrow \sqrt{z_1^2 + z_2^2 + z_3^2 + \cdots + z_{n-1}^2} \leq z_n$ 意思是，最后的分量大于前面的分量的平方和的开方（欧氏定理?）

一个例子：$\sqrt{x^2 + y^2} \leq z \Leftrightarrow x^2 + y^2 \leq z^2, z \geq 0$



(2-扩展)SOCP 的一种等价写法 equicalent SOCP formulation

$$
\begin{aligned}
&\min c^T x \\
&\text{s.t. } Ax - b \geq_{socp} 0
\end{aligned}
$$

$$
\Rightarrow (A, b) = \begin{pmatrix} A^0, b^0 \\ A_n, b_n \end{pmatrix}
$$

$$
\Rightarrow \begin{array}{ll} \min & c^T x \\ \text{s.t.} & \left\| A^0 x - b^0 \right\| \leq A_n x - b_n \end{array}
$$

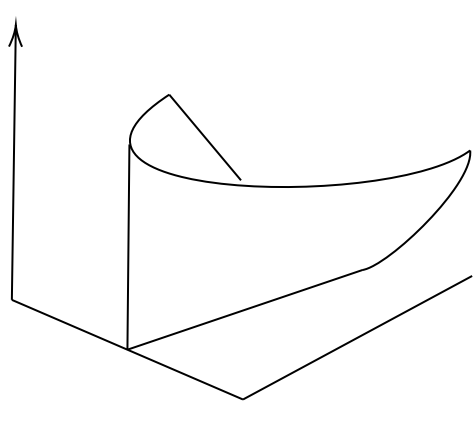
$$
z = (z_1, \cdots, z_{n-1}, z_n)^T \Rightarrow z = (z^0, z_{n-1}, z_n)^T \Rightarrow \sqrt{z_1^2 + z_2^2 + z_3^2 + \cdots + z_{n-1}^2} \leq z_n \Leftrightarrow \left\| z^0 \right\| \leq z_n
$$

(3)Semidefinite Programming(SDP) 半正定规划

$$
\begin{aligned}
&\min c^T x \\
&\text{s.t. } x_1 F_1 + x_2 F_2 + \cdots + x_n F_n - G \geq_{socp} 0
\end{aligned}
$$

上述的目标函数是线性的，约束条件要求 left hand side 为半正定矩阵；$F_1, F_2, \cdots, F_n, G$ are symmetrical matrices(所以下面的例子是对称的，y 和 y). $A - B \geq_{SDP} 0$ indicates that $A - B$ is a semidefinite matrix.

$$
S = \left\{ (x, y, z) : \begin{pmatrix} x & y \\ y & z \end{pmatrix} \geq_{SDP} 0 \right\}
$$



Note: SDP 的对偶还是 SDP 锥

(4) 上述问题的逻辑

$$LP \in SOCP \in SDP()$$

- Interior point methods can be applied to SCOP and SDP in a similar way to LP.
- Many problems can be casted as SOCP or SDP problems.
- Free software CVX

1.2 Applications in Finance

1.2.1 Mean-Variance model 均值-方差模型

1.2.2 Mean-CVaR model 均值-CVaR 模型

Varicance is a risk measure.（在投资组合里常用）

Value-at-Risk（VaR，风险值）：是凸的，α- percentile of distribution of random variable (quantile based) (a smallest value such that probability that random variable less than or equal to this value is greater than or equal to α)

Conditional Value-at-Risk（CVaR）=conditional expectation that the loss is greater than or equal to VaR (moment-quantile based)

$$\mathrm{CVaR}_\alpha = E\left[\xi \mid \xi \geq \mathrm{VaR}_\alpha(\xi)\right]$$

CVaR 是一个一致性风险度量。CVaR minization is a convex program, but VaR may not. CVaR is a coherent risk measure.

CVaR 的优化问题

x 是决策变量，ξ 是与损失相关的随机变量。ζ 是引进来的一个变量。

1.2.3 Dynamic financial planning 动态财务规划 (Stochastic programming approach)

最小方差前沿：如果两个资产的话，那么在 $\sigma - E(r)$ 的面上，最小方差前沿是双曲线；如果是多个资产的话，那么有效集的边界是双曲线，有效集是包括了双曲线里的内容。

分离定理的本质：风险资产的投资组合与风险厌恶态度无关。

分离定理对资产组合选择的启示：

- 由分离定理，资产组合选择问题可以分为两个独立的工作，即资本配置决策（Capital allocation decision）和资产选择决策（Asset allocation decision）
- 资本配置决策：考虑资金在无风险资产和风险资产组合之间的分配。
- 资产选择决策：在众多的风险证券中选择适当的风险资产构成资产组合
- 由分离定理，基金公司可以不必考虑投资者偏好的情况下，确定最优的风险组合。

资产组合理论的优点：

- 首次对风险和收益进行精确的描述，解决对风险的衡量问题，使投资学从一个艺术迈向科学。
- 分散投资的合理性为基金管理提供理论依据。单个资产的风险并不重要，重要的是组合的风险。
- 从单个证券的分析，转向组合的分析。

资产组合理论的缺点

- 当证券的数量较多时，计算量非常大，使模型应用受到限制。
- 解的不稳定性。(根据历史数据总是不稳定的)
- 马科维茨及其学生夏普于是寻求更为简便的方法，并发展出了定价理论，这就是后面要学习的指数模型和 CAPM。

1.2.4 Proof for CAPM by using Fakas lemma, (参考：link)

引理 1.1 (Farkas' lemma)**.** *Let* $\mathbf{A} \in \mathbb{R}^{m\times n}$ *and* $\mathbf{b} \in \mathbb{R}^m$*. Then exactly one of the following two statements is true:*

(1) *There exists* $\mathbf{x} \in \mathbb{R}^n$ *such that* $\mathbf{Ax} = \mathbf{b}$ *and* $\mathbf{x} \geq 0$*.*

(2) *There exists* $\mathbf{y} \in \mathbb{R}^m$ *such that* $\mathbf{A}^T\mathbf{y} \geq 0$ *and* $\mathbf{b}^T\mathbf{y} < 0$*.*

Here, $\mathbf{x} \geq 0$ *means* $x_i \geq 0$ *for all* i*.*

定义 1.1 (Arrow-Debreu Security)**.** Assume a two-period world: today and tomorrow, with tomorrow having S possible states. An asset that pays 1 in state s and 0 otherwise is called an Arrow-Debreu (A-D) security. Its price is denoted by q_s.

For an arbitrary asset i with payoff $X_{i,s}$ in state s, its price is:

$$P_i = \sum_s q_s X_{i,s} \tag{3}$$

In a complete market with no arbitrage, Stiemke's lemma implies that the equation $\mathbf{Xq} = \mathbf{P}$ has a solution $\mathbf{q} > 0$. Rewriting equation (3):

$$P_i = \sum_s q_s X_{i,s}$$

This states that the price of asset i equals the sum of its state payoffs times the state prices. No arbitrage implies a positive state-price vector, sometimes referred to as the fundamental theorem of asset pricing.

Implications Multiplying Equation (3) by the probability π_s of state s:

$$P_i = \sum_s \pi_s \frac{q_s}{\pi_s} X_{i,s} \tag{4}$$

Define $M = \frac{q_s}{\pi_s}$, known as the stochastic discount factor (SDF). Then:

$$P_i = \sum_s \pi_s M X_{i,s} \iff P_i = \mathbb{E}[M X_{i,s}] \tag{5}$$

Dividing both sides by P_i yields the fundamental equation of asset pricing:

$$1 = \mathbb{E}[M R_i]$$

For a risk-free asset with return R_{rf} (non-stochastic):

$$1 = \mathbb{E}[MR_{rf}], \Rightarrow 1 = R_{rf}\mathbb{E}[M], \Rightarrow \frac{1}{\mathbb{E}[M]} = R_{rf}$$

Thus, the risk-free rate equals the inverse of the expected SDF.

Asset pricing model From equation (5): $P = \mathbb{E}[MX]$. Using the identity $\mathrm{Cov}(M, X) = \mathbb{E}[MX] - \mathbb{E}[M]\mathbb{E}[X]$:

$$P = \mathrm{Cov}(M, X) + \mathbb{E}[M]\mathbb{E}[X], \Leftrightarrow P = \frac{1}{R_{rf}}\mathbb{E}[X] + \mathrm{Cov}(M, X) \tag{6}$$

Rewriting Equation(6) and dividing both sides by P:

$$\mathbb{E}[X] = PR_{rf} - R_{rf}\mathrm{Cov}(M, X) \tag{7}$$

$$\mathbb{E}[R] = R_{rf} - R_{rf}\mathrm{Cov}(M, R) \tag{8}$$

Equation (8) indicates that the expected return of any asset depends on the risk-free rate and its covariance with the SDF. Expressing this in terms of excess return:

$$\mathbb{E}[R] - R_{rf} = -R_{rf}\mathrm{Cov}(M, R) \tag{9}$$

Furthermore, multiplying and dividing by $Var(M)$ and substituting back $R_{rf} = \frac{1}{\mathbb{E}[M]}$,

$$\mathbb{E}[R] - R_{rf} = \frac{\mathrm{Cov}(M, R)}{\mathrm{Var}(M)} \cdot \left(-\frac{\mathrm{Var}(M)}{\mathbb{E}[M]}\right)$$

Equation (9) is called a beta pricing model and sometimes written as

$$\mathbb{E}[R] - R_{rf} = \beta_{M,R} \cdot \lambda_M,$$

where $\beta_{M,R}$ is the quantity of risk in the asset. Note that λ_M does not directly depend on the individual return but on the first and second moments of the SDF.

The theory of SDF relates several independent asset pricing models, some commonly known examples are:

- In the CAPM, the SDF is a linear combination of the market portfolio.
- In consumption (micro-founded) models, the SDF is the marginal rate of substitution of utility today and in the future.
- In Arrow-Debreu pricing, the SDF is a function of aggregate demand.

推荐阅读：

- Fama E F, French K R. The capital asset pricing model: Theory and evidence[J]. Journal of Economic Perspectives, 2004, 18(3): 25-46.

2 收益-风险理论

$$r_i = E(r_i) + \varepsilon_i$$

$$r_i = E(r_i) + m + e_i$$

其中 m 是系统的，e_i 是非独立的。

2.1 CAPM 资本资产定价模型

资本资产定价模型 (Capital Asset Pricing Model, CAPM) 是现代金融学的奠基石。该模型对于资产风险与其期望收益率之间的关系给出了精确的预测。CAPM 是基于风险资产期望收益均衡基础上的预测模型。

- 它提供了一种估计潜在投资项目收益率的方法。
- 模型使得我们能对不在市场交易的资产同样做出合理的估价。

William Sharpe（1964）, John Lintner（1965）与 Jan Mossin（1966）等做出了非常重要的贡献。

2.1.1 资本资产定价模型

基本内容

- 完全竞争市场
- 投资周期相同，短视行为
- 投资者可以在固定的无风险利率基础上借入或贷出任何额度的资产
- 投资环境无摩擦
- 投资者符合 Markowitz 理性
- 同质预期（投资者关于有价证券收益率的概率分布是一致的）

市场组合的风险溢价: $E(r_M) - r_f = \bar{A}\sigma_M^2$

单个证券的风险溢价：$E\left(r_i\right) - r_f = \frac{\mathrm{Cov}(r_i, r_M)}{\sigma_M^2}\left[E\left(r_M\right) - r_f\right] = \beta_i\left[E\left(r_M\right) - r_f\right]$

这个单个证券的风险溢价，实际上是在马科维茨最优投资组合时的一阶最优性条件。

投资者对市场组合的选择

根据假设 2、3、5、6，投资者的最优风险资产组合相同 → 若某一个股票未包含在最优资产组合中，会怎样?

2.1.2 CAPM 的严格证明

资本市场线 CML $\bar{r}_p = r_f + \frac{\bar{r}_m - r_f}{\sigma_m}\sigma_p$ 注意 $\bar{r}$ 指的是期望。

命题 2.1. 切点组合中任一资产 i 的期望收益满足

证明. 考虑持有权重 w 的资产 i，权重 $1-w$ 的市场组合 m 构成一个新的资产组合，由组合计算公式有：

$$\bar{r_w} = w\bar{r_i} + (1-w)\bar{r_m}$$

$$\sigma_w = \sqrt{w^2\sigma_i^2 + (1-w)^2\sigma_m^2 + 2w(1-w)\sigma_{im}}$$

运用参数化求导的方法，也即：

$$\frac{d\bar{r}_w}{dw} = \bar{r}_i - \bar{r}_m$$

$$\frac{d\sigma_w}{dw} = \frac{w\sigma_i^2 + (w-1)\sigma_m^2 + (1-2w)\sigma_{im}}{\sigma_w}$$

两个比一下

$$\left.\frac{d\bar{r}_w}{d\sigma_w}\right|_{w=0} = \left.\frac{d\bar{r}_w/dw}{d\sigma_w/dw}\right|_{w=0} = \frac{(\bar{r}_i - \bar{r}_m)\,\sigma_m}{\sigma_{im} - \sigma_m^2}$$

得到 r 与 σ 的导数。这叫参数化求导。

该斜率与资本市场线相等，则：

$$\frac{(\bar{r}_i - \bar{r}_m)\,\sigma_m}{\sigma_{im} - \sigma_m^2} = \frac{\bar{r}_m - r_f}{\sigma_m}$$

则有：

$$\overline{r_i} = r_f + \frac{\sigma_{im}}{\sigma_m^2}\left(\bar{r}_m - r_f\right) = r_f + \beta_i\left(\bar{r}_m - r_f\right)$$

也就有 $r_i = \bar{r}_i + \varepsilon_i = r_f + \beta_i\left(\bar{r}_m - r_f\right) + \varepsilon_i$。与另一式（市场模型）$r_i - r_f = \alpha_i + \beta_i\left(r_m - r_f\right) + \varepsilon_i$ 相比，有何区别?

→ 一个是 CAPM，一个是物理模型，虽然都是夏普搞出来的东西。

首先，两个 ε_i 是不同的。上面的 $\bar{r_m}$ 是均值（期望），不确定性都在 ε_i 里。而下边的 r_m 是随机变量，因此下面的 ε_i 只包含了非系统的部分，系统的部分在 r_m 里包括了。

但是，这两个 β 从计算的角度上是一样的。CAPM 中的 β_i 是定义的，其被定义为 $\frac{Cov(r_i, r_m)}{\sigma_m^2}$。而下边市场模型中的 β_i 可求解为：

$$\frac{Cov(r_i, r_m)}{\sigma_m^2} = \frac{Cov\left[(r_f + \alpha_i + \beta_i(r_m - r_f) + \varepsilon_i), r_m\right]}{\sigma_m^2} = \frac{Cov\left[(\beta_i r_m + \varepsilon_i), r_m\right]}{\sigma_m^2} = \beta_i\frac{Cov(r_m, r_m)}{\sigma_m^2} = \beta_i$$

证券的系统风险本质上是该证券与市场上所有证券的协方差加权和。$r_m=\sum_{j=1}^{n} w_j r_j$：

$$\beta_i=\frac{\sigma_{im}}{\sigma_m^2}=\frac{cov(r_m,r_i)}{\sigma_m^2}=\frac{\sum_{j=1}^{n} w_j cov(r_j,r_i)}{\sigma_m^2}$$

一般地，由于一种证券不可能与市场上所有证券之间都相互独立，故系统风险不为 0。

CAPM 的一个简单推导：

假设有 n 个风险资产，$r_1,\cdots,r_n$ 和 r_f。我们的目标是

$$\min \frac{\delta}{2}w^T\Sigma w-(\bar{\mathbf{r}}^{\mathbf{T}}w+w_0 r_f) s.t. \sum_{i=1}^{n} w_i+w_0=1$$

也即

$$\min \frac{\delta}{2}w^T\Sigma w-(\bar{r}-\mathbf{1}r_f)^T w$$

注意 $\sum_{i=1}^{n}=w_i r_i+w_i r_f$

因为是凸优化问题，其 F.O.C 是 $\delta\Sigma w_i-(\bar{r}-\mathbf{1}r_f)=0$，也即第一行是

$$\delta\left[cov(r_i,r_1)w_1+\cdots+cov(r_i,r_n)w_n\right]=\bar{r_1}-r_f$$

$$\delta cov(r_1,w_1r_1+w_2r_2+\cdots+w_nr_n)=\bar{r_1}-r_f$$

因为

$$\delta=\frac{E(r_m)-r_f}{\sigma_m^2}$$

（这是由于 $y=\frac{E(r_m)-r_f}{\delta\sigma_m^2}$,$\delta$ 实际上跟风险厌恶程度相关），所以上式改写为

$$\frac{E(r_m)-r_f}{\sigma_m^2}Cov(r_i,r_m)=\bar{r_1}-r_f$$

CAPM 是终点，而不是起点。

2.2 套利定价理论（APT）

如果资产暴露在同样的风险下，那么他们的收益应当相同。

因素组合：各因素的 β 对自己因素的敏感性为 1，对其他因素的敏感性为 0。

组合构建方法：按比例 β_{p1}、β_{p2} 、$(1-\beta_{p1}-\beta_{p2})$ 投资于因素组合 1、因素组合 2、无风险资产。已知 $r_p=E(r_p)+\beta_{p1}F_1+\beta_{p2}F_2$。由于

$$r_{F_1}=E(r_{F_1})+1*F_1+0*F_2,\quad r_{F_2}=E(r_{F_2})+0*F_1+1*F_2$$

，则有

$$r_c = \beta_{p_1} E(r_{F_1}) + \beta_{p_2} E(r_{F_2}) + r_f - \beta_{p_1} r_f - \beta_{p_2} r_f + \beta_{p_1} F_1 + \beta_{p_2} F_2$$

，则由无套利，

$$\begin{aligned} E(r_p) &= \beta_{p_1} E(r_{F_1}) + \beta_{p_2} E(r_{F_2}) + r_f - \beta_{p_1} r_f - \beta_{p_2} r_f \\ &= r_f + \beta_{p_1} \left[E(r_{F_1}) - r_f) \right] + \beta_{p_2} \left[(E(r_{F_2}) - r_f) \right] \end{aligned} \tag{10}$$

多因素套利定价理论推导

投资者都相信证券 i 的收益受 m 个共同因素影响，如下：

$$r_i = E(r_i) + \beta_{i1} F_1 + \beta_{i2} F_2 + \cdots + \beta_{im} F_m + e_i$$

套利组合必须同时具备如下三个特征：

(1) 它是一个不需要投资者任何额外资金的组合

$$w_1 + w_2 + w_3 + \cdots + w_n = 0$$

(2) 套利组合对任何因素都没有敏感性

$$w_1 \beta_{1j} + w_2 \beta_{2j} + \cdots + w_n \beta_{nj} = 0, j = 1, \cdots, m$$

(3) 套利组合的预期收益率必须是正值

$$w_1 E(r_1) + w_2 E(r_2) + \cdots + w_n E(r_n) > 0$$

考虑优化问题（线性规划）：

$$\max_w E(r_1) w_1 + \cdots + E(r_n) w_n$$

$$s.t. w_1 + w_2 + w_3 + \cdots + w_n = 0$$

$$\beta_{11} w_1 + \beta_{21} w_2 + \cdots + \beta_{n1} w_n = 0$$

$$\vdots$$

$$\beta_{1m} w_1 + \beta_{2m} w_2 + \cdots + \beta_{nm} w_n = 0$$

一个可行解：$w = 0$。是否有非 0 解使得目标值大于 0 呢？

如果有 w^* 可以，那么 kw^* 也可以。那么这个目标函数可以随意地大。因此若目标值大于 0，则其是有无界解（则其没有最优解）。因此，无套利等价于有最优解等价于满足最优性条件。

上述线性规划问题最优性条件 (K-T 条件，充分必要条件):

$$
\begin{aligned}
&\min_{w} -E\left(r_1\right) w_1-\cdots-E\left(r_n\right) w_n \\
&s.t. w_1+w_2+w_3+\cdots+w_n=0 \\
&\beta_{11} w_1+\beta_{21} w_2+\cdots+\beta_{n 1} w_n=0 \\
&\vdots \\
&\beta_{1 m} w_1+\beta_{2 m} w_2+\cdots+\beta_{n m} w_n=0
\end{aligned}
$$

直接得其最优性条件为：$-\nabla f(w)+\sum \lambda_j \nabla h_j(w)=0$ 也即：

$$
-\begin{bmatrix} E(r_1) \\ E(r_2) \\ \cdots \\ E(r_n) \end{bmatrix}+\lambda_0\begin{bmatrix} 1 \\ 1 \\ \cdots \\ 1 \end{bmatrix}+\sum_{j=1}^m \lambda_j\begin{bmatrix} \beta_{1j} \\ \beta_{2j} \\ \cdots \\ \beta_{nj} \end{bmatrix}=0 \tag{11}
$$

存在 λ_k 使得 $E(r_i)=\lambda_0+\beta_{i1}\lambda_1+\beta_{i2}\lambda_2+\cdots+\beta_{im}\lambda_m$，其中 λ_k 为投资者承担一个单位 k 因素风险的补偿额，风险的大小由 β_{ik} 表示。进一步有：

$$
\lambda_0=r_f
$$

$$
\lambda_j=E(r_{Fj})-r_f, j=1,\cdots,m
$$

其实就是影子价格。让所有的 β 为 0，则 $\lambda_0=r_f$。让除了某个 β 之外的 β 为 0，则 $\lambda_j=E(r_{Fj})-r_f, j=1,\cdots,m$。

因此有：

$$
E(r_i)=r_f+\beta_{i1}\left[E(r_{F1}-r_f)\right]+\cdots+\beta_{im}\left[E(r_{Fm}-r_f)\right]
$$

2.3 有效市场假说 EMH

略

2.4 组合风险分析

2.4.1 金融风险度量 VaR

VaR 的基本概念: VaR (Value-at-Risk, 风险值)：金融行业的风险测度标准 (巴塞尔协议)。

VaR 的通俗定义：在给定时期内和给定置信概率水平下，投资组合可能面临的最大损失（的下限）。

P(某段时期内某投资价值 ≤ VaR) = 1-c, 其中 c 是置信水平（90 95 99%）

设投资组合 W 在未来给定时间段内价值的概率密度函数为 f(w)，给定置信水平 c (90/95/99%)，VaR 定义为：

- 绝对 VaR= W^*
- 相对 VaR= $E(W) - W^*$

其中 W^* 满足：

$$c = \int_{W^*}^{\infty} f(w)dw$$

VaR 的计算

- 非参数法: 历史数据，核估计
- 半参数法：用了参数，但不依赖于对分布的假设
- 参数法：参数 + 假设分布是什么

VaR 参数估计

设投资组合的收益率为 R，初始财富为 W_0，期末财富为：$W = W0(1+R)$。令 $W^* = W(1+R^*)$ 得

$$\begin{aligned} 1-c = P\{W \leq W^*\} = P\{W \leq W_0(1+R^*)\} = P\left\{\frac{W}{W_0} \leq 1+R^*\right\} \\ = P\{1+R \leq 1+R^*\} = P\{R \leq R^*\} = \int_{-\infty}^{R^*} f(r)dr \end{aligned} \tag{12}$$

因此，绝对 VaR 的计算可由 W 的 1-c 的百分位数换算为 R 的 1-c 的百分位数（W 的随机性源于 R，R 的分位数就是 W 的分位数）。

$$W^* = W_0(1+R^*)$$

相对 VaR: $VaR = E(W) - W^* = W_0(\mu - R^*)$ 其中 $\mu = E(R)$。

正态分布下的 VaR

如果是正态分布的话，就可以做到 不受分布参数的影响。这样可以做投资组合的凸优化；如果不是正态分布，那就是非凸的，不好做。当然，如果仅仅是测度的话可以。

2.5 投资组合风险分解

$$\mathrm{VaR}_p = \alpha\sigma_p W_0 = \alpha\sqrt{w' \sum w} W_0$$

一种风险管理方法：分解协方差矩阵（判断其对什么因素敏感）对风险的暴露。

一般而言 VaR 是满足次可加性的（组合的风险小于风险的组合），但有时候也不一定。

3 债券组合管理

3.1 利率期限结构和利率风险

- 马科维茨的资产组合理论（均值 - 方差模型）
- 夏普等的 CAPM
- 罗斯的 APT

上述三种现代投资学基本理论在关于证券的风险与收益关系的分析中，将证券视为一个高度抽象的概念，本篇将证券具体化，以一种简单而基本证券——债券为例，对特定的证券投资理论进行论述。

关注：利率的变化（利率期限结构）；债券信用评估。也就是利率风险和信用风险

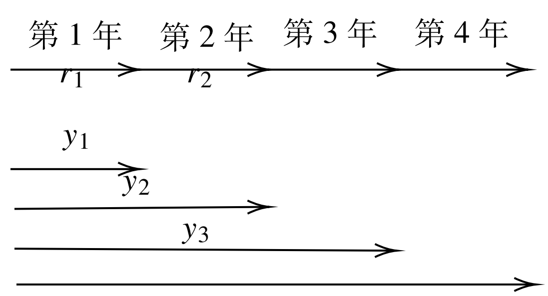
- 利率的期限结构理论：债券期限与利率水平的关系。
- 债券的久期与凸性理论：含义、计算方法及在债券投资管理中的运用。

3.1.1 利率的期限结构

利率的期限结构 (term structure of interest rates)

- 确定的利率期限结构
- 不确定的利率期限结构
- 利率期限结构理论

探讨债券的期限长度与利率水平的关系。



短期利率： 给定期限的利率 r。零息票债券定价：

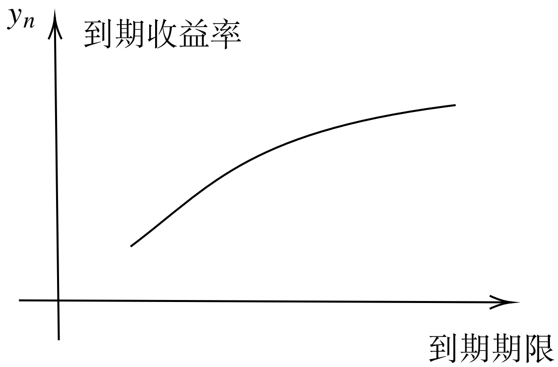
$$PV = \frac{Par}{(1+r_1)(1+r_2)\cdots(1+r_n)}$$

到期收益率：

$$PV = \frac{Par}{(1+y_n)^n}$$

零息票债券的到期收益率也称即期利率。

收益率曲线（yield curve）：反映不同到期时间的债券的到期收益率与到期时间关系的曲线。



发行含息证券时，用求出的到期收益率进行折现。零息票债券构成的收益率曲线可以用于对其他债券的定价。

$$P = \frac{C_1}{1+y_1} + \frac{C_2}{(1+y_2)^2} + \cdots + \frac{C_n}{(1+y_T)^T}$$

$$P = \frac{C_1}{1+r_1} + \frac{C_2}{(1+r_1)(1+r_2)} + \cdots + \frac{C_n}{(1+r_1)\cdots(1+r_T)}$$

如果现实中不存在这些零息票债券，可以用很多债券的数据对 $\frac{1}{1+y_1}$、$\frac{1}{1+y_2}$ 等进行估计。

远期利率

$$(1+y_n)^n = (1+y_{n-1})^{n-1}(1+r_n)$$

$$1+r_n = \frac{(1+y_n)^n}{(1+y_{n-1})^{n-1}}$$

如果将远期利率定义为 f_n，则有

$$1+f_n = \frac{(1+y_n)^n}{(1+y_{n-1})^{n-1}}$$

远期利率与未来实际短期利率不一定相等。只有在利率事前确定的条件下，远期利率才一定等于未来短期利率。

如果投资者购买两年期的债券并持有至到期，那么其收益：$(1+r_1)(1+f_2) = (1+y_2)^2$。如果只是投资一年，第二年年初卖掉，那么到手的收益为：$\frac{(1+y_2)^2}{1+r_2}$。r_2 对于现在而言是未知的。所以求个期望也就是：$E(\frac{(1+y_2)^2}{1+r_2})$

如果是持有一年到期的债券，收益是确定的，为 $1+r_1$。所以，如果要求**有短期偏好的投资者**不持有短期而持有长期再卖掉，要求：

$$E(\frac{(1+y_2)^2}{1+r_2}) > 1+r_1$$

$$\frac{(1+y_2)^2}{1+E(r_2)} > 1+r_1$$
$$\frac{(1+r_1)(1+f_2)}{1+E(r_2)} > 1+r_1$$
$$\frac{(1+f_2)}{1+E(r_2)} > 1$$
$$f_2 > E(r_2)$$

远期利率要大于未来预期的短期利率。f_2 实际上是投资者对未来短期利率预期的上界。

反之，如果要求**有长期偏好的投资者**不持有长期而持有短期，则有

$$E(r_2) > f_2$$

远期利率要小于未来预期的短期利率。f_2 实际上是投资者对未来短期利率预期的下界。

结论

- 如果投资者偏好短期投资，就要求远期利率 f 大于期望的短期利率 r；
- 如果投资者偏好长期投资，则要求期望的短期利率 r 大于远期利率 f。
- 远期利率是否等于未来期望的短期利率取决于投资者对利率风险的承受情况，也取决于他们对债券期限长短的偏好。**注意前述结论都是在投资者风险厌恶的前提下进行讨论的**

利率期限结构理论

期限结构理论是指说明长短期债券利率水平的关系的理论。

- 预期假定 (expectations hypothesis)：预期假定是最简单的期限结构理论。这一理论认为远期利率等于市场整体对未来短期利率的预期。
- 流动偏好 (liquidity preference)：投资者有不同的期限偏好，有些偏好短期债券，有些偏好长期债券。要求远期利率与期望的未来短期利率之间有一个溢价。大多数人偏好短期投资。

$$\frac{(1+y_2)^2}{1+y_1} = 1+f_2$$

因为流动性溢价，$f_2 > r_1$

$$(1+f_2)^2 > (1+y_2)^2 = (1+f_2)(1+r_1) > (1+r_1)^2$$

所以照理来说，这个线是往上走的，但永远小于 f。

$$1+f_3 > 1+y_3 > (1+y_3)^{3/2} = [(1+f_3)(1+y_2)]^{1/2} > 1+y_2$$

3.2 债券久期与凸性理论

决定利率风险的因素：债券定价规则证实了决定利率风险有以下因素：

- 期限长度
- 息票利率
- 当前到期收益率（初始 YTM）

3.2.1 债券久期

久期是把债券每次利息或本金的支付时间进行加权平均所得到的期限。因此，久期测度的是债券的实际持有期限。或者说，是债券支付的未来现金流（本息）的到期期限的加权平均值，也称为债券的平均期限，它是债券的有效期限。

- 零息票债券：由于期间没有支付息票利息，债券的实际持有期限就是债券的到期期限 (duration is equal to maturity for zero coupon bonds)。
- 息票债券：由于债券到期之前，每期都会支付息票利息，从而使债券的实际期限缩短。

$$P = \frac{C_1}{1+y_1} + \frac{C_2}{(1+y_2)^2} + \cdots + \frac{C_n}{(1+y_T)^T}$$

令 $y = y_1 = y_2 = \cdots = y_T$ （实际上是假设了水平的到期收益率曲线），则有：

$$P = \frac{C_1}{1+y} + \frac{C_2}{(1+y)^2} + \cdots + \frac{C_n}{(1+y)^T}$$

这个到期收益率，仅仅是对这一债券而言，不一定通用于其他债券。我们真正的定价要用不同期限的零息票债券的到期收益率作为折现因子才可以。

$$\begin{aligned} \frac{dP}{P} &\approx \frac{1}{P} \cdot \frac{dP}{dy} dy + \frac{1}{P} \cdot \frac{1}{2} \frac{d^2P}{dy^2} (dy)^2 \\ &= -D^* dy + \frac{1}{2} c (dy)^2 \end{aligned} \tag{13}$$

$\frac{dP}{P}$ 表明的是价格变化的百分比，也即修正久期乘以利率变化的百分比取负。久期越大，对利率的变化越敏感。利率变化引起债券价格实际上升的幅度比久期的线性估计要高，而下降的幅度要小。在其他条件相同时，人们应该偏好凸性大的债券。凸性大，YTM 上升，债券价格下跌不明显；凸性大，YTM 下降，债券价格上升不明显。

3.2.2 基于 N-S 利率模型的凸性免疫

N-S 曲线来估计收益率曲线，是比较常用的。the Nelson-Siegel model of yield to maturity can be represented as:

$$r(m) = \beta_0 + (\beta_1 + \beta_2)\frac{1 - e^{-\frac{m}{\tau}}}{\frac{m}{\tau}} - \beta_2 e^{-\frac{m}{\tau}}$$

where m is 到期日。

将利率曲线（利率风险）的变化转换为 β_0、β_1、β_2 的变化。

3.3 信用债券组合管理——CreditMetrics 方法

债券的评级情况：用信用评级转移矩阵去刻画；

债券违约后：用债券回收率去刻画。

知道上述信息还不够，要知道债券的联合转移状态是什么。那么这就需要用到债券的相关性信息。债券信用级别的联合变化可以等价于公司价值的联合变化；进一步地，可以简化为公司股票价格的联合变化。

3.3.1 信用风险的基本概念

信用风险 (Credit Risk): 由于对方未能偿还其债务而造成的损失的风险。

信用风险管理的重要性比市场风险影响大。

信用风险度量的基本要素

- 信用等级
- 违约率
- 回收率

3.3.2 CreditMetrics

略

4 衍生资产定价与风险计算

4.1 期权组合管理

注意，如果基础资产价格为负，那么就不能用那个欧式期权定价公式，因为负数不能取对数了。

期权报价：报隐含波动率。

好处：（1）可以直接计算价格（2）可以反映人们的预期

B-S 公式作用：

（1）报价（工作语言）

（2）风险管理

如：

$$f_0 = g(S_0)$$

$$f_t = g(S_t) = g(S_0) + g'(S_0)\Delta S_t$$

$$\Delta f_t = g'(S_0)\Delta S_t$$

$$\alpha\Delta S_t + \beta\Delta f_t = \alpha\Delta S_t + \beta g'(S_0)\Delta S_t$$

（3）定价

4.1.1 B-S 公式的推导

首先我们要知道：

几何布朗运动：

$$S_{t+\Delta t} = S_t e^{\left(\left(\mu-\frac{1}{2}\sigma^2\right)\Delta t+\sigma\varepsilon_{N(0,1)}\sqrt{\Delta t}\right)}$$

其中 ε 服从 $N(0,1)$ 的标准正态分布。

具体地，

$$\widetilde{\varepsilon} = a + b\varepsilon\, N(a, b^2)$$

其中 $a = \left(\mu - \frac{1}{2}\sigma^2\right)$, $b^2 = \sigma^2\Delta t$

$$\ln\frac{S_{t+\Delta t}}{S_t} \sim N(a\Delta t, b^2\Delta t)$$

为什么可以假设上式服从正态分布？

$$\ln\left(\frac{S_T}{S_0}\right) = \ln\left(\frac{S_T}{S_{T-1}}\frac{S_{T-1}}{S_{T-2}}\cdots\frac{S_1}{S_0}\right) = \ln\left(\frac{S_T}{S_{T-1}}\right) + \ln\left(\frac{S_{T-1}}{S_{T-2}}\right) + \cdots + \ln\left(\frac{S_1}{S_0}\right)$$

上述各项其实都是随机变量。假设他们是独立同分布的，(独立就可以，甚至不需要同分布）且 $\ln\left(\frac{S_1}{S_0}\right) \sim N(\theta, \gamma^2)$。那么

$$\ln\left(\frac{S_T}{S_0}\right) \sim N(T\theta, T\gamma^2)$$

因此，期权定价公式中，出现了正态分布，且运动的均值和方差与 T 相关。

实践中，月度收益率的正态性比较好。如果是日度，可能出现尖峰厚尾的现象，即频率比较高的数据，正态性会差一点，因为不一定是独立的。指数数据的话，正态性会比个股更好一些。

4.1.2 随机过程相关

随机过程三大内容：泊松过程、马氏过程、布朗运动。

给定时间，它是一个随机变量；从时间轴上看，它是一个随机过程。

基本维纳过程：

定义： 一个随机过程 W_t，在微小时间间隔 Δt 之间的变化为 ΔW，如果

1.$W_0 = 0$

2. 瞬时增量为 $\Delta W = \varepsilon$ 即

$$\varepsilon \sim N(0,1) \Rightarrow \Delta W \sim N(0, \sqrt{\Delta t})$$

3. 在任意在两个微小时间段内的改变量 ΔW 是独立的

则称 W_t 的运动遵循 Wiener (维纳) 过程或布朗运动。

性质：

Wiener 过程 (长时间段内) 的增量。将 $[0, T]$ 按等距划成 N 等分

$$\Delta W_T = W_T - W_0 = \sum_{i=1}^{N} \varepsilon_i \sqrt{\Delta t}$$

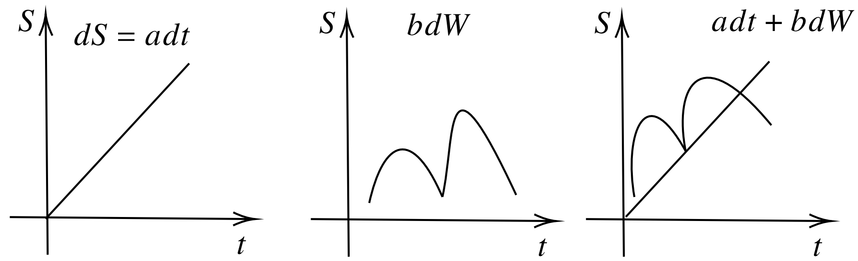
$$\frac{T}{\Delta t} = N$$

$$\Delta W_T \sim N(0, \sqrt{T})$$

可以看出，ΔW_T 的方差只依赖于时间差距 T。称

$$dW = adt + b\varepsilon\sqrt{dt}\varepsilon\sqrt{dt} \quad (\Delta t \to 0)$$

为标准维纳过程或称为标准布朗运动。



漂移率：单位时间内变量 W 的均值。

方差率：单位时间内变量 W 的方差。

一般维纳过程：称

$$dS = adt + bdW = adt + b\varepsilon\sqrt{dt}$$

为一般 Wiener 过程。其中 dW 为标准维纳过程。

漂移率（扩散系数）：a

方差率（漂移系数）：b^2

另：

$$\Delta S = a\Delta t + b\Delta W = a\Delta t + b\varepsilon\sqrt{\Delta t}$$

4.1.3 Ito 引理

一般情况下允许 a 和 b 都和随机变量 S 有关，即被称为为 Ito (伊藤) 过程的随机微分方程 SDE。

X 是 Ito 过程，如果

$$dX = a(X,t)dt + b(X,t)dW$$

即 a, b 都是 X 与 t 的函数。

X_t 是一个维纳过程 W_t 驱动的 Ito 过程，即

$$dX = \mu(X,t)dt + \sigma(X,t)dW$$

$f(X,t)$ 是 X 与 t 的函数。$f(X,t)$ 随时间的变化 $df(X,t)$ 必须用 (随机) 泰勒公式来描述。

应用：期权定价（期权是股票的函数）

具体地：

$f(X,y)$ 如果可微，那么

$$df(X,y) = f_x(x,y)dx + f_y(x,y)dy$$

但，如果 X 满足

$$dX = \mu(X,t)dt + \sigma(X,t)dW$$

f 是 X 与 t 的函数，则

$$\Delta(X,t) = f_X^0 \Delta X + f_t^0 \Delta t + \frac{1}{2} f_{XX}^0 (\Delta X)^2 + o(\Delta t)$$

把 $(\Delta X)^2$ 代入就可以了，就有

$$df = \left(\frac{\partial f}{\partial X}\mu + \frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial X^2}\sigma^2\right)dt + \frac{\partial f}{\partial X}\sigma dW$$

那么 f 也是一个 Ito 过程。上述其实只是记号。Ito 引理其实是从积分的角度去推的，因为随机变量没有导数。对于函数而言，其极限的定义为：$lim \|\xi_n - \xi\| = 0$ 对于随机变量而言，其极限的定义其实是在均方分析的角度下定义的。$E(\xi_n - \xi)^2 = 0$，其实也就是依概率收敛到 0。

$(dW)^2 = dt$ 其实也是在均方分析的意义下相等。

4.2 期权 VaR 的计算

假设基础资产的变化 $\dfrac{\Delta S}{S_0} \sim N(\theta, \gamma^2)$(风险因子服从正态分布)
这个假设其实就简化了很多问题了。
衍生资产（比如期权）$f(S)$ 呢?
一阶的话：

$$f(S) \approx f(S_0) + f'(S_0)\Delta S$$

$$\Delta f = f'(S_0)\Delta S$$

二阶的话：

$$f(S) \approx f(S_0) + f'(S_0)\Delta S + \frac{1}{2} f''(S_0)(\Delta S)^2$$

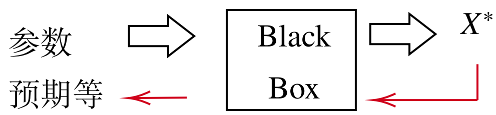
$$\Delta f = f'(S_0)\Delta S + \frac{1}{2} f''(S_0)(\Delta S)^2$$

期权的市场组合是 0（多空头），所以不能用 CAPM 那一套去定价。

5 Black-Litterman 模型

反推市场投资者对当前的预期。

人类简史：天气是一个一级混沌市场，但金融市场不是。我的预期不会影响明天的天气，但我的预期会影响金融市场。



如：CAPM

$$\max \mu - \frac{1}{2}\delta W^T \Sigma W$$

最优性条件：

$$\mu - \frac{\delta}{2}\Sigma w^* = 0$$

w^* 是市场组合 (已知)。因此，我们可以通过

$$\mu = \frac{\delta}{2}\Sigma w^*$$

反推 μ。结合投资者的行为（优化），知道投资者的结果（市场组合），可以反推投资者的预期。

当然，也可以写成

$$\mu = \delta \Sigma w^*$$

当 $\delta = \frac{E(r_M)-r_f}{\sigma_M^2}$，这就是 CAPM。