FINANCE 1~3 RECAP

1. Risk Management

Coherent Risk Measure: A risk measure $\rho(X)$ defined on losses is said to be coherent if it satisfies the following properties.

- I. $\rho(X+c) = \rho(X) + c$ Translation Invariance
- II. $\rho(\lambda X) = \lambda \rho(X), \lambda > 0$ Positive Homogeneity (not affected by the units of measurement)
- III.$\rho(X+Y) \leq \rho(X) + \rho(Y)$Subadditivity
- IV.If $\mathbb{P}(X \leq Y) = 1$then $\rho(X) \leq \rho(Y)$ Monotonicity

VaR is defined by

$$VaR_\alpha(X) = \inf\{x|\mathbb{P}(L > x) \leq 1-\alpha\} = \inf\{x|G(x) \geq \alpha\}$$

$= G^{-1}(\alpha)$ when it exists.

When the distribution function is continuous,

$$CVaR_\alpha(X) = \mathbb{E}[X|X \geq VaR_\alpha(X)]$$

otherwise

$$CVaR_\alpha(X) = \frac{\mathbb{E}\left[X \cdot \mathbf{1}_{\{X \geq VaR_\alpha(X)\}}\right]}{1-\alpha} + VaR_\alpha(X) \cdot (1 - \alpha - \mathbb{P}(X \geq VaR_\alpha(X)))$$

The following table shows the properties of variance, VaR and CVaR (Expected Shortfall, Conditional Tail Expectation)

	σ	VaR	CVaR
I	No	Yes	Yes
II	Yes	Yes	Yes
III	Yes	No	Yes
IV	No	Yes	Yes

Normal Distribution:

$$f(x) = \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \cdot \frac{1}{\sqrt{2\pi\sigma^2}}$$

and VaR and CVaR for normal random variable are given by

$$VaR_\alpha = \mu + \sigma N^{-1}(\alpha)$$

and

$$CVaR_\alpha = \mu + \frac{\sigma}{1-\alpha} N\left(\frac{VaR_\alpha - \mu}{\sigma}\right)$$

Rating Agencies: Moody's and S&P

Internal Rating: All banks use to assess the creditworthiness of borrowers and counterparties.

Altman's Z-Score: used to estimate default probabilities based on accounting data

$$Z = 1.2X_1 + 1.4X_2 + 3.3X_3 + 0.6X_4 + 0.999X_5$$

where X_1 is the ratio of Working Capital to Total Assets, X_2 is the ratio of Retained Earnings to Total Assets, X_3 is the ratio of Earnings Before Interests and Taxes (EBIT) to Book Value of Total Liabilities and X_5 is the ratio of Sales to Total Assets.

Hazard Rates means Default Intensities. Survival function $S(t) = \Pr(\tau > t) = 1 - F(t)$ and the hazard rate $\lambda(t) = -\frac{S'(t)}{S(t)}$ and it can be easily derived that

$$S(t) = \exp\left(-\int_0^t \lambda(\nu)\, d\nu\right)$$

and Reduced Form models assume λ follows a SDE.

The **Average Hazard Rate** $\bar{\lambda}(t)$ is defined by $\bar{\lambda}(t) = -\frac{\ln(S(t))}{t}$ and $S(t) = \exp\left(-\bar{\lambda}(t)\cdot t\right)$ and $F(t) = 1 - \exp\left(-\bar{\lambda}(t)\cdot t\right)$.

Cumulative Hazard Rate is defined by $\Gamma(t) = -\ln(S(t))$ and Hazard Rate can also be written as

$$\lambda(t) = \frac{f(t)}{1 - F(t)} = \lim_{h\to 0} \mathbb{P}(\tau \le t + h | \tau > t)$$

and CDF in terms of Hazard Rate can be represented as

$$F(t) = 1 - \exp(-\Gamma(t)) = 1 - \exp\left(-\int_0^t \lambda(\nu)\, d\nu\right)$$

Hazard Rate can be interpreted as default intensities by observing the following derivation

$$\begin{aligned}\mathbb{P}(t < \tau \le t + h | \tau > t) &= \frac{S(t) - S(t+h)}{S(t)} \\ &= 1 - \exp\left(-\int_t^{t+h} \lambda(s)\, ds\right) \\ &\approx h \cdot \lambda(t)\end{aligned}$$

Recovery Rate RR is used to indicate that when an obligor defaults on a loan, the percentage of the value of the contract remains after default.

Loss give default LGD is defined by $LGD = 1 - RR$.

Moody's has found the following equation provided a reasonable fit to historical experience (1983-2004)

$$\overline{RR} = 0.52 - 6.9 \times Average\, Default\, Rate$$

Consider a zero-coupon bond, the implied default probability can be derived as

$$q = \frac{1 - e^{-sT}}{1 - RR}$$

by equating the theoretical bond price $e^{-(r+s)T}$ and the market price $e^{-rT}(q \cdot RR + 1 - q)$.

The difference between implied default probability q and real world default probability p: q is derived under risk neutral probability measure for pricing purpose. For real world risk management and portfolio selection, p derived from accounting data should be used.

Structural Models: Merton, First Passage (Firm defaults at $\tau = \inf\{t | A_t \le b\}$ and $\mathbb{P}[t \le \tau \le t + dt] = -\frac{\ln(b/A_0)}{\sqrt{2\pi\sigma^2 t^3}} \exp\left(\frac{-\left(\ln(b/A_0) - \left(\mu - \frac{\sigma^2}{2}\right)t\right)^2}{2t\sigma^2}\right) dt$), Boundary Crossing.

Several Spreads:

- TED Spread: is a measure of credit risk for inter-bank lending. It is the difference between 1) the risk free 3 month U.S. treasury bill rate; and 2) the three month LIBOR rate, which represents the rate at which banks typically lend to each other.

Merton's Model:

- Capital Structure: Liability L due at T, Asset A_t follows a Geometric Brownian Motion
$$dA_t = \mu A_t dt + \sigma A_t dZ_t$$
hence by Itô's Lemma
$$A_t = A_0 \exp\left(\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma \cdot Z_t\right)$$
- At T, what happens
  - If $A_t \geq L$, bondholder receives L, shareholders receive $A_T - L$
  - If $A_t < L$, bondholder receives A_t,shareholders receive nothing.
- Payoff function for
  - Shareholder: Call option on the assets $\max(A_t - L, 0)$
  - Bondholder: $A_t - \max(A_t - L, 0)$
- Use Black-Scholes to price the equity of the firm
$$E_t = A_t \Phi(d_1) - e^{-rT} L \Phi(d_2)$$
where $d_1 = \frac{\ln\left(\frac{A_t}{L}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}$ and $d_2 = d_1 - \sigma\sqrt{T-t} = \frac{\ln\left(\frac{A_t}{L}\right) + \left(r - \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}$
- Probability of default in Merton's Model
  - Real world: $\mathbb{P}[A_T \leq L] = \mathcal{N}\left(\frac{\ln\left(\frac{L}{A_0}\right) - \left(\mu - \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}\right)$
  - Risk neutral: $\mathbb{Q}[A_T \leq L] = \mathcal{N}\left(\left(\frac{\ln\left(\frac{L}{A_0}\right) - \left(r - \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}\right)\right)$
- Default happens $A_T < L$ is equivalent to $Z_T < \frac{\ln\left(\frac{L}{A_0}\right) - \left(\mu - \frac{\sigma^2}{2}\right)T}{\sigma}$, i.e., $Z < \mathcal{N}^{-1}(PD)$ where Z is a standard normal random variable.

Portfolio Credit Risk:

CAPM:
$$R_i = r_f + \beta_i (R_M - r_f) + \epsilon_i$$
CAPM:: Security Market Line
$$\mathbb{E}[R_i] = r_f + \beta_i (\mathbb{E}[R_M] - r_f)$$
where $\beta_i = \frac{cov(R_i, R_M)}{\sigma_M^2}$ and β_i is the slope of the regression line from fitting the equation ($\beta = 1$: Neutral,$\beta >$ 1:Aggressive,$\beta <$ 1:Defensive Stocks)
$$R = \alpha + \beta R_M + \epsilon$$
Note: Securities that lie below the SML are overpriced and vice versa.

Single Factor Model: (Developed by Vasicek, similar in thinking to CAPM) also called single factor Gaussian copula model. The credit index corresponding to a given name as a systematic component plus an idiosyncratic component
$$Y_n = \beta_n Z + \sqrt{1 - \beta_n^2} \cdot \epsilon_n$$
where Z and ϵ_n are i.i.d. standard normal and the default indicator has the form
$$\mathbf{1}_{\{Y_n \leq H_n\}}$$
Clearly it is a Bernoulli RV with mean PD_n and variance $PD_n \cdot (1 - PD_n)$.

- The correlation between the credit indices of two different names is

$$\rho_{Y_n,Y_m} = \frac{cov\left(Y_n, Y_m\right)}{\sigma_{Y_n}\sigma_{Y_m}} = \mathbb{E}\left[Y_n Y_m\right] - \mathbb{E}\left[Y_n\right]\mathbb{E}\left[Y_m\right] = \beta_n\beta_m\mathbb{E}\left[Z^2\right] = \beta_n\beta_m$$

- The correlation between the default indicators of two different names is

$$\rho = \frac{\mathbb{E}\left[\mathbf{1}_{\{Y_n \le H_n\}} \cdot \mathbf{1}_{\{Y_m \le H_m\}}\right] - PD_n \cdot PD_m}{\sqrt{PD_n \cdot \left(1 - PD_n\right) \cdot PD_m \cdot \left(1 - PD_m\right)}}$$

- Bivariate normal distribution with correlation r is given by

$$\mathcal{N}_2\left(x, y, r\right) = \mathbb{P}\left(X \le x, Y \le y\right) = \int_{-\infty}^{x}\int_{-\infty}^{y} e^{-\frac{x^2 - 2rxy + y^2}{2\left(1-r^2\right)}} \frac{dxdy}{2\pi\sqrt{1-r^2}}$$

- Total portfolio loss is given by

$$L_T = \sum_{n=1}^{N} w_n\left(1 - RR_n\right) \cdot \mathbf{1}_{\{Y_n \le H_n\}}$$

and the systematic loss is given by

$$L_S = \mathbb{E}\left[L_T | Z\right]$$

which is often referred to as the "large portfolio limit".

- Conditional expectation for a given name is deduced as follows

$$\begin{aligned}
\mathbb{E}\left[\mathbf{1}_{\{Y_n \le H_n\}}|Z\right] &= \mathbb{P}\left[Y_n \le H_n | Z\right] \\
&= \mathbb{P}\left[\beta_n Z + \sqrt{1-\beta_n^2} \cdot \epsilon_n \le H_n | Z\right] \\
&= \mathbb{P}\left[\left[\epsilon_n \le \frac{H_n - \beta_n Z}{\sqrt{1-\beta_n^2}} | Z\right]\right] \\
&= \mathcal{N}\left(\frac{\mathcal{N}^{-1}\left(PD_n\right) - \beta_n Z}{\sqrt{1-\beta_n^2}}\right)
\end{aligned}$$

Hence the systematic loss is a decreasing function of the global systematic factor Z by observing

$$L_S = \sum_{n=1}^{N} w_n\left(1 - RR_n\right)\mathcal{N}\left(\frac{\mathcal{N}^{-1}\left(PD_n\right) - \beta_n Z}{\sqrt{1-\beta_n^2}}\right)$$

- The VaR $L_{S,\alpha}$ is simply given by

$$L_{S,\alpha} = \sum_{n=1}^{N} w_n\left(1 - RR_n\right)\mathcal{N}\left(\frac{\mathcal{N}^{-1}\left(PD_n\right) - \beta_n Z_{1-\alpha}}{\sqrt{1-\beta_n^2}}\right)$$

by applying **Euler's Theorem**

$$L_{S,\alpha} = \sum_{n=1}^{N} w_n \frac{\partial L_{S,\alpha}}{\partial w_n}$$

where

$$\frac{\partial L_{S,\alpha}}{\partial w_n} = \left(1 - RR_n\right)\mathcal{N}\left(\frac{\mathcal{N}^{-1}\left(PD_n\right) - \beta_n Z_{1-\alpha}}{\sqrt{1-\beta_n^2}}\right)$$

Basel Accord II: The single factor model forms the basis for the credit risk approach of the second Basel Capital Accord. The economic capital is based on VaR which is at the 99.9% confidence level with a one year time horizon.

The capital charge for an instrument is

$$EAD \cdot LGD \cdot (WCDR - PD) \cdot MA$$

where

$$WCDR = \mathcal{N}\left(\frac{\mathcal{N}^{-1}(PD_n) - \beta_n Z_{0.001}}{\sqrt{1-\beta_n^2}}\right)$$

Multi-Factor Model: Credit index is defined by

$$Y_n = \sum_{k=1}^{K} \beta_{n,k} Z_k + \sqrt{1 - \sum_{k=1}^{K} \beta_{n,k}^2} \cdot \epsilon_n$$

and the total portfolio loss is given by

$$L_T = \sum_{n=1}^{N} w_n (1 - RR_n) \cdot \mathbf{1}_{\{Y_n \le H_n\}}$$

and the systematic loss as before is given by

$$L_S = \mathbb{E}[L_T | Z_1, \cdots, Z_K] = \sum_{n=1}^{N} w_n (1 - RR_n) \mathcal{N}\left(\frac{\mathcal{N}^{-1}(PD_n) - \beta_n \sum_{k=1}^{K} \beta_{n,k} Z_k}{\sqrt{1 - \sum_{k=1}^{K} \beta_{n,k}^2}}\right)$$

Note that in the multi-factor model, VaR can not be computed explicitly unless all factor correlations are the same. Moreover, Large Portfolio Limit says

$$\lim_{N\to\infty} L_N - \mathbb{E}[L_N | Z] = 0$$

Industry Examples: **Credit Metrics/KMV**

$$Y = \beta \cdot Z + \epsilon$$

where Y is the vector of default indicators, Z is the vector of systematic factors, β is the matrix of factor loadings and ϵ is the vector of idiosyncratic factors.

Normal Mixture Models: Default indicators are given by

$$Y = m(W) + \sqrt{W} \cdot X$$

where $X = \beta \cdot Z + \epsilon$, $Z \sim N(0, \Sigma), \epsilon \sim N(0, I)$. A key example is when v/W has a chi-squared distribution with v degrees of freedom, this is a multivariate t model.

Bernoulli Mixture Models: Default indicators are given by

$$y = (y_1, \cdots, y_N) \in \{0, 1\}^N$$

and

$$\mathbb{P}[Y = y | Z = z] = \Pi_{n=1}^{N} PD_n(Z)^{y_n} (1 - PD_n(Z))^{1-y_n}$$

Exchangeable Bernoulli Mixtures: Conditional Default Probability $Q = PD(Z)$ and the distribution of number of defaults is given by

$$P(M = n) = \begin{pmatrix} N \\ n \end{pmatrix} \int_0^1 q^n (1-q)^{N-n} \, dG(q)$$

and there are three variations on

- Beta: $Q \sim B(a, b)$
- Probit: $Q = \Phi(\mu + \sigma Z), Z \sim \mathcal{N}(0, 1)$
- Logit: $Q = F(\mu + \sigma Z), Z \sim \mathcal{N}(0, 1), F(x) = (1 + \exp(-x))^{-1}$

Copulas: is used to model multivariate random variables. The basic idea is to separate the modelling into Marginal Distributions and Dependence Structure(copula)

Fact: Suppose that X is a random variable with CDF $F(x)$, then $U = F(X)$ has a uniform distribution on $(0, 1)$. This can be seen from discussion below.

$$\mathbb{P}(U \leq x) = \mathbb{P}(F(X) \leq x) = \mathbb{P}\left(X \leq F^{-1}(x)\right) = F\left(F^{-1}(x)\right) = x$$

hence any random variable X with CDF $F(x)$ can be generated via generating a uniform random variable U on $(0, 1)$ followed by applying $F^{-1}(\cdot)$ to this random variable.

Side note: If one wants to generate normal random variables, firstly Cholesky decompose $\Sigma = LL^T$, then simulate a normal random vector with mean zero and variance-covariance matrix I, lastly define your $X = \mu + LZ$

Correlation is not a good indicator of dependence since it only measures linear dependence.

- Formally, a copula is the joint distribution function of N random variables with uniform $(0, 1)$ marginal distributions.
- If $X = (X_1, \cdots, X_N)$ is a random vector with joint distribution function F, and marginal distributions $F_1, \cdots, F_N$ then the copula of X is

$$C(u_1, \cdots, u_N) = F\left(F_1^{-1}(u_1), \cdots, F_N^{-1}(u_N)\right)$$

- Alternatively, the copula of X is the joint distribution of $(F(X_1), \cdots, F(X_N))$

Suppose we want to generate N random variables with a given copula C, and marginal distributions $G_1, \cdots, G_N$

- First simulate $U_1, \cdots, U_N$ with joint distribution C
- Then compute $X_n = G_n^{-1}(U_n)$

Example: Gaussian Copula

Suppose we want to have N random variables with the same copula as an N dimensional normal random variables with variance-covariance matrix Σ, but with marginal distributions $G_1, \cdots, G_N$

First simulate $X = (X_1, \cdots, X_N)$ from the distribution $N(0, \Sigma)$, then define $Y_n = G_n^{-1}(N(X_n))$

Coefficient of Tail Dependence: Let X_1 and X_2 be random variables with marginal distribution functions F_1 and F_2. Then the coefficient of upper tail dependence of X_1 and X_2 is given by

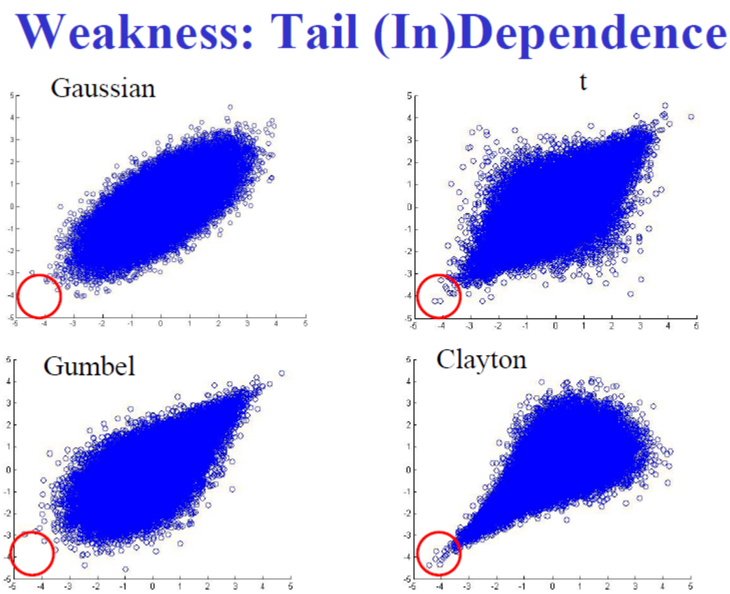
$$\lambda_u = \lambda_u(X_1, X_2) := \lim_{q \to 1^-} \mathbb{P}\left(X_2 > F_2^{-1}(q) \,|X_1 > F_1^{-1}(q)\right)$$

notice that Λ_u depends only on the copula of (X_1, X_2), not on the marginal distributions.

For Gaussian Copula, $\lambda_u = 0$ asymptotically. For the copula with a t distribution, $\lambda_u = 2t_{v+1}\left(-\sqrt{\frac{(v+1)(1-\rho)}{1+\rho}}\right)$

Archimedean Copulas

- General Form: $C(u_1, \cdots, u_N) = \Phi^{-1}\left(\sum_{n=1}^N \phi(u_n)\right)$
- Clayton: $\phi(t) = \theta^{-1}\left(t^{-\theta} - 1\right), \theta \geq -1$
- Gumbel: $\phi(t) = (-\ln t)^{\theta}, \theta \geq 1$
- Frank: $\phi(t) = -\ln\left(\frac{e^{-\theta t}-1}{e^{-\theta}-1}\right)$



2. Robust Portfolio Optimization

2.1. **Black-Litterman Model.** Incorporate investor views into the mean-variance model. Base model makes an equilibrium assumption, i.e., investors who do not have a view on the market will hold the market portfolio. In order to achieve the above goal, one has to

- Step One: Define Market Equilibrium
- Step Two: Express Investor Views
- Step Three: Combine and Estimate

2.1.1. *CAPM Assumptions.*

$$\mathbb{E}\left[R_i\right] - R_f = \beta_i \left(\mathbb{E}\left[R_M\right] - R_f\right)$$

where $\beta_i = \frac{cov(R_i, R_M)}{\sigma_M^2}$, $R_M = \sum_{n=1}^{N} w_n^b R_n$. Thus

$$\mathbb{E}\left[R_i\right] - R_f = \frac{\mathbb{E}\left[R_M\right] - R_f}{\sigma_M^2} \cdot \sum_{n=1}^{N} w_n^b cov\left(R_i, R_n\right)$$

For simplicity, CAPM assumptions are represented as

$$\Pi = \begin{pmatrix} \mathbb{E}\left[R_1\right] - R_f \\ \vdots \\ \mathbb{E}\left[R_N\right] - R_f \end{pmatrix} = \delta \cdot \Sigma w^b$$

where scalar $\delta = \frac{\mathbb{E}[R_M] - R_f}{\sigma_M^2}$, Σ is an $n \times n$ variance-covariance matrix with the i-th row, j-th column component being $cov\left(R_i, R_j\right)$ and w^b is an n-by-1 vector containing the weights.

Uncertainty in Mean Estimation is modeled as

$$\Pi = \mu + \epsilon_\Pi, \epsilon_\Pi \sim N\left(0, \tau \cdot \Sigma\right)$$

where τ is chosen to represent the degree of uncertainty in the estimate of the vector of expected returns.

2.1.2. *Investor Views.*

$$q = P\mu + \epsilon_q, \epsilon_q \sim N(0, \Omega)$$

where q is a K dimensional random vector (K views), P is a $K \times N$ matrix expressing the views and Ω is a $K \times K$ covariance matrix expressing the degree of confidence in the views. For example, $N = 5$ stocks

- Stock 1 will have a return of 1.5%
- Stock 3 will outperform Stock 2 by 4%.

Above two views can be incorporated into the following system

$$\begin{pmatrix} 1.5\% \\ 4\% \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \\ \mu_5 \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \end{pmatrix}$$

2.1.3. *Combine and Estimate.* A linear model for μ combining equilibrium and investor views is given by

$$y = X\mu + \epsilon, \epsilon \sim N(0, V)$$

where

$$y = \begin{pmatrix} \Pi \\ q \end{pmatrix}, X = \begin{pmatrix} I \\ P \end{pmatrix}, V = \begin{pmatrix} \tau\Sigma & 0 \\ 0 & \Omega \end{pmatrix}$$

The generalized lease squares estimator for the above model is given by

$$\begin{aligned}
\hat{\mu}_{BL} &= \left(X'V^{-1}X\right)^{-1} X'V^{-1}y \\
&= \left(\begin{pmatrix} I & P' \end{pmatrix} \begin{pmatrix} (\tau\Sigma)^{-1} & \\ & \Omega^{-1} \end{pmatrix} \begin{pmatrix} I \\ P \end{pmatrix}\right)^{-1} \begin{pmatrix} I & P' \end{pmatrix} \begin{pmatrix} (\tau\Sigma)^{-1} & \\ & \Omega^{-1} \end{pmatrix} \begin{pmatrix} \Pi \\ q \end{pmatrix} \\
&= \left(\begin{pmatrix} I & P' \end{pmatrix} \begin{pmatrix} (\tau\Sigma)^{-1} \\ \Omega^{-1} \end{pmatrix}\right)^{-1} \begin{pmatrix} I & P' \end{pmatrix} \begin{pmatrix} (\tau\Sigma)^{-1}\Pi \\ \Omega^{-1}q \end{pmatrix} \\
&= \left((\tau\Sigma)^{-1} + P'\Omega^{-1}P\right)^{-1} \left((\tau\Sigma)^{-1}\Pi + P'\Omega^{-1}q\right)
\end{aligned}$$

Note that if we only consider views, $\hat{\mu}_I = (P'P)^{-1} P'q$ and $P(P'P)^{-1} P = I$. Therefore after decomposition of above, we get

$$\begin{aligned}
\hat{\mu}_{BL} &= \underbrace{\left((\tau\Sigma)^{-1} + P'\Omega^{-1}P\right)^{-1} (\tau\Sigma)^{-1}\Pi}_{Equlibrium(CAPM)} + \underbrace{\left((\tau\Sigma)^{-1} + P'\Omega^{-1}P\right)^{-1} P'\Omega^{-1}P\hat{\mu}_I}_{Views} \\
&= w_\Pi \Pi + w_q \hat{\mu}_I
\end{aligned}$$

where $w_\Pi + w_q = I$. We can further get

$$\hat{\mu}_{BL} = \underbrace{\Pi}_{CAPM} + \underbrace{\tau \cdot \Sigma P' \left(\Omega + \tau \cdot P\Sigma P'\right)^{-1} (q - \Pi)}_{Tilt}$$

and the variance of the Black-Litterman estimator can be calculated explicitly as

$$var\left(\hat{\mu}_{BL}\right) = \left(X'V^{-1}X\right)^{-1} = \left((\tau \cdot \Sigma)^{-1} + P'\Omega^{-1}P\right)^{-1}$$

2.2. **Two Fund Theorem Recap.** Any mean variance efficient portfolio can be written as a combination of two particular portfolios.

The mean variance optimization problem is given by

$$\begin{array}{rl} & \min_w \frac{1}{2} w^T \Sigma w - \tau w^T \mu \\ st. & w^T \mathbf{1} = 1 \end{array}$$

The Lagrangian can be written as

$$L(w, \lambda) = \frac{1}{2} w^T \Sigma w - \tau w^T \mu - \lambda \left(w^T \mathbf{1} - 1 \right)$$

leading to the optimality condition

$$\Sigma w - \tau \mu - \lambda \mathbf{1} = \mathbf{0}$$

and the optimal portfolio is given by

$$w = \tau \Sigma^{-1} \mu + \lambda \Sigma^{-1} \mathbf{1}$$

and this is a linear combination of the minimum risk portfolio $w_{mr} = \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}^T \Sigma^{-1} \mathbf{1}}$ and the market portfolio $w_m = \frac{\Sigma^{-1}\mu}{\mathbf{1}^T \Sigma^{-1} \mathbf{1}}$.

Suppose that there is a risk-free asset with return R_0 and N risky assets with return $R_n, n \in \{1, \cdots, N\}$. The utility maximization problem of investor k is given by

$$\begin{array}{rl} & \max \mathbb{E} \left[U \left(\sum_{n=0}^{N} w_n R_n \right) \right] \\ st. & \sum_{n=1}^{N} w_n = 1 \end{array}$$

Suppose that the utility function for investor k satisfies

$$U_k'(x) = (A_k + B_k x)^{-c}$$

where A_k, B_k can vary between investors, but c is the same for all investors. Prove the following:

Theorem 1. *(Two-Fund Theorem) All investors will hold a combination of the risk-free security and a fixed portfolio of the risky assets, i.e., while investors may place different amounts of wealth in the risky assets, they will always invest in the risky assets in the same proportions.*

Proof. Omitted here. □

2.3. **Can the sharp ratio between two risky assets exceed the slope of capital market line (CML)?.** Why CAPM? What is the simple proof of this market equilibrium model?

We know that a marginal utility of asset i is equal to the expected return minus its contribution to the volatility of the portfolio, i.e.,

$$\mathbb{E}[r_i] - a\sigma_{i,m} = k$$

where a is the risk aversion, $\sigma_{i,m} = cov(r_i, r_m)$ and k must be the same for every asset because of the optimality assumptions (Otherwise one would invest more on the other asset with higher risk premium). Notice that

$$\mathbb{E}[r_f] - a\sigma_{f,m} = r_f = k$$

and

$$\mathbb{E}[r_m] - a\sigma_{m,m} = r_f$$

which is equivalent to

$$a = \frac{\mathbb{E}[r_m] - r_f}{var(r_m)}$$

therefore the CAPM model can be represented as

$$\mathbb{E}\left[r_i\right]-r_f=\frac{\sigma_{i,m}}{var\left(r_m\right)}\left(\mathbb{E}\left[r_m\right]-r_f\right)$$

The expected payoff of the portfolio is given by

$$R=\sum_{n=1}^{N}w_nR_n$$

and the variance of R is trivially derived as

$$\begin{aligned}
var\left(R\right) &= \mathbb{E}\left[\left(R-\mathbb{E}\left[R\right]\right)^2\right]\\
&= \mathbb{E}\left[R^2\right]-\mathbb{E}\left[R\right]^2\\
&= \mathbb{E}\left[\left(\sum_{n=1}^{N}w_nR_n\right)^2\right]-\mathbb{E}\left[\sum_{n=1}^{N}w_nR_n\right]^2\\
&= \mathbb{E}\left[\sum_{i=1}^{N}\sum_{j=1}^{N}w_iw_jR_iR_j\right]-\sum_{i=1}^{N}\sum_{j=1}^{N}w_iw_j\mathbb{E}\left[R_i\right]\mathbb{E}\left[R_j\right]\\
&= \sum_{i=1}^{N}\sum_{j=1}^{N}w_iw_j\left(\mathbb{E}\left[R_iR_j\right]-\mathbb{E}\left[R_i\right]\mathbb{E}\left[R_j\right]\right)\\
&= \sum_{i=1}^{N}\sum_{j=1}^{N}w_iw_jcov\left(R_i,R_j\right)\\
&= w^T\Sigma w
\end{aligned}$$

where Σ is the covariance matrix with (i,j)-th component being $cov\left(R_i,R_j\right)$. For the purpose of illustration, let $w_i=\frac{1}{N},\forall i\in\{1,\cdots,N\}$, $cov\left(R_i,R_j\right)=a,\forall i,j$and $var\left(R_i\right)=b,\forall i$, we can see that

$$var\left(R\right)=\frac{1}{N}a+\frac{N\left(N-1\right)}{N^2}b=\frac{1}{N}a+\frac{N-1}{N}b$$

the covariance can not be diversified away as we increase the number of assets in the portfolio.

The problem can be mathematically stated as is it possible to have

$$\frac{\mathbb{E}\left[R_i\right]-\mathbb{E}\left[R_j\right]}{\sigma\left(R_i-R_j\right)}>\frac{\mathbb{E}\left[R_{mv}\right]-r_f}{\sigma\left(R_{mv}\right)}$$

The short answer is NO since the above implies

$$\beta_i-\beta_j>\frac{\sigma\left(R_i-R_j\right)}{\sigma\left(R_{mv}\right)}$$

hence

$$\frac{cov\left(R_i-R_j,R_{mv}\right)}{\sigma\left(R_i-R_j\right)\sigma\left(R_{mv}\right)}>1$$

which is not possible.

3. Numerical Computation for Financial Modelling

It is always good to know that by definition

$$Var\left(Y|X\right)=\mathbb{E}\left[\left(Y-\mathbb{E}\left[Y|X\right]\right)^{2}|X\right]$$

and

$$Var\left(Y\right)=\mathbb{E}\left[var\left(Y|X\right)\right]+var\left(\mathbb{E}\left[Y|X\right]\right)$$

3.1. Brownian Motion.

The Brownian motion with drift is given by

$$dX=\alpha dt+\sigma dZ$$

where αdt is the drift term, σ is the volatility and dZ is a random term in the form of $dZ=\phi\sqrt{dt}$, in which ϕ is a standard normal random variable. Notice that

$$\mathbb{E}\left[dX\right]=\alpha dt$$

and

$$Var\left(dX\right)=\sigma^{2}dt$$

and it can be both numerically or analytically verified that, as $dt\to0$, with probability one, $dZ^{2}\to dt$. In another word, dZ^{2} becomes non-stochastic (Quadratic Variation).

Lemma 2. ***(Itô's Lemma)*** *Suppose we have some function $G=G\left(S,t\right)$ and $dS=a\left(S,t\right)dt+b\left(S,t\right)dZ$, then*

$$dG=\left(aG_{S}+G_{SS}\frac{b^{2}}{2}+G_{t}\right)dt+G_{S}bdZ$$

Proof. We take a Taylor expansion on dG up to the second order approximation

$$\begin{aligned}dG &= G_{S}dS+G_{t}dt+\frac{1}{2}G_{SS}\left(dS\right)^{2}+\frac{1}{2}G_{tt}\left(dt\right)^{2}+G_{S,t}\left(dS\cdot dt\right)\\ &= G_{S}\left(adt+bdZ\right)+G_{t}dt+\frac{1}{2}G_{SS}b^{2}dt\\ &= \left(aG_{S}+\frac{b^{2}}{2}G_{SS}+G_{t}\right)dt+bG_{S}dZ\end{aligned}$$

□

Application of Itô's Lemma: Solve GBM $dA_{t}=\mu A_{t}dt+\sigma A_{t}dZ_{t}$

Notice that

$$d\ln A_{t}=\left(\mu A_{t}\frac{1}{A_{t}}+\frac{\sigma^{2}A_{t}^{2}}{2}\left(-\frac{1}{A_{t}^{2}}\right)\right)dt+\sigma A_{t}\frac{1}{A_{t}}dZ_{t}=\left(\mu-\frac{\sigma^{2}}{2}\right)dt+\sigma dZ_{t}$$

Taking integration on both side from 0 to t, one can easily get

$$\ln\frac{A_{t}}{A_{0}}=\left(\mu-\frac{\sigma^{2}}{2}\right)t+\sigma\left(Z_{t}-Z_{0}\right)$$

hence

$$A_{t}=A_{0}\exp\left(\left(\mu-\frac{\sigma^{2}}{2}\right)t+\sigma Z_{t}\right)$$

$(Z_{0}=0)$ by the definition of Wiener Process.

3.2. **Black-Scholes Equation.**

- The stock price follows GBM. ($dS_t = \mu S_t dt + \sigma S_t dZ_t$, i.e., $S_t = S_0 e^{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma Z_t}$)
- The risk free rate of return is a constant r. ($dB_t = rB_t dt$, i.e., $B_t = B_0 e^{rt}$)
- There are no arbitrage opportunities
- Short selling is permitted

Let $C(S,t)$ be the price at time t of the derivative. Premium is $C(S_0, 0)$ and payoff at maturity T is given by $C(S_T, T) = (S_T - K)^+$ for a call.

By Itô's formula,

$$dC = \frac{\partial C}{\partial t} dt + \frac{\partial C}{\partial S} dS + \frac{1}{2} \frac{\partial^2 C}{\partial S^2} (dS)^2$$

therefore

$$dC = \left(C_t + \mu S C_S + \frac{1}{2} \sigma^2 S^2 C_{SS} \right) dt + C_S \sigma S dZ$$

By no arbitrage pricing theorem, price of the derivative is equal to the price of a replicating portfolio.

We use V_t to denote the value of the portfolio at time t. It is simply composed of stocks and bond, i.e., $V_t = a_t S_t + b_t B_t$. Assume it is a self-financing strategy.

$$V_t = V_0 + \int_0^t a_s dS_t + \int_0^t b_s dB_s$$

then

$$\begin{aligned} dV_t &= a_t dS_t + b_t dB_t \\ &= (a_t \mu S + b_t r B_t) dt + a_t \sigma S dZ_t \end{aligned}$$

Since V is a replicating portfolio, i.e., $V_t = C(S_t, t), \forall t \in [0, T]$, we get $dV_t = dC$. Thus

$$a_t = C_S$$

and

$$C_t + \mu S C_S + \frac{1}{2} \sigma^2 S^2 C_{SS} = a_t \mu S + b_t r B_t$$

Rearrange the above terms and simplify in use of $b_t = \frac{C - C_S S_t}{B_t}$, we get

$$-rC + C_t + rSC_S + \frac{1}{2} \sigma^2 S^2 C_{SS} = 0$$

which is the classic Black-Scholes equation.

Feynman-Kac Formula: (Under Risk Neutral Measure)

If S follows $dS = rSdt + \sigma S dZ$ and C satisfies $-rC + C_t + rSC_S + \frac{1}{2}\sigma^2 S^2 C_{SS} = 0$, $C(S_T, T)$ gives payoff at T, then

$$C(S_t, t) = \mathbb{E}\left[e^{-r(T-t)} C(S_T, T) | \mathcal{F}_t \right]$$

The premium for a European call is

$$
\begin{aligned}
C(S_0,0) &= \mathbb{E}\left[e^{-rT}C(S_T,T)\,|\mathcal{F}_0\right] \\
&= e^{-rT}\mathbb{E}\left[\max\left(S_0e^{\left(r-\frac{\sigma^2}{2}\right)T+Z_T}-K,0\right)\right] \\
&= e^{-rT}\mathbb{E}\left[\left(S_0e^{\left(r-\frac{\sigma^2}{2}\right)T+Z_T}-K\right)\cdot\mathbf{1}_{\{S_T>K\}}\right] \\
&= e^{-rT}S_0e^{\left(r-\frac{\sigma^2}{2}\right)T}\mathbb{E}\left[e^{Z_T}\cdot 1_{\{S_T>K\}}\right]-e^{-rT}K\cdot\mathbb{E}\left[\mathbf{1}_{\{S_T>K\}}\right]
\end{aligned}
$$

Notice that

$$
\mathbb{E}\left[\mathbf{1}_{\{S_T>K\}}\right] = \mathbb{P}(S_T>K)=\mathbb{P}(\ln S_T>\ln K)
$$

Since $\ln S_T\sim N\left(\ln S_0+\left(r-\frac{\sigma^2}{2}\right)T,\sigma^2T\right)$, we can get

$$
\begin{aligned}
& \mathbb{P}(\ln S_T>\ln K) \\
=\ & \mathbb{P}\left(\frac{\ln S_T-\ln S_0-\left(r-\frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}>\frac{\ln K-\ln S_0-\left(r-\frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}\right) \\
=\ & 1-\mathcal{N}\left(\frac{\ln\frac{K}{S_0}-\left(r-\frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}\right) \\
=\ & \mathcal{N}\left(\frac{\ln\left(\frac{S_0}{K}\right)+\left(r-\frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}\right)
\end{aligned}
$$

The first term can be also derived in a similar fashion. Actually one can use $\mathbb{E}\left[e^X\cdot 1_{\{e^X>a\}}\right]=e^{m+\frac{V}{2}}\Phi\left(\frac{m+V-\ln a}{\sqrt{V}}\right)$ where $X\sim N(m,V)$ or use change of measure method.

In all, the pricing formula for Black-Scholes is given by

$$
C(S_0,0)=S_0\mathcal{N}\left(\frac{\ln\left(\frac{S_0}{K}\right)+\left(r+\frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}\right)-Ke^{-rT}\mathcal{N}\left(\frac{\ln\left(\frac{S_0}{K}\right)+\left(r-\frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}\right)
$$

3.3. **Discrete Hedging.** Bear in mind these Greeks for Black-Scholes Model

Table 6.1: Formulæ for European call.

	Call
Payoff	$\max(S-K,0)$
Value V Black–Scholes value	$Se^{-D(T-t)}N(d_1)-Ke^{-r(T-t)}N(d_2)$
Delta $\frac{\partial V}{\partial S}$ Sensitivity to underlying	$e^{-D(T-t)}N(d_1)$
Gamma $\frac{\partial^2 V}{\partial S^2}$ Sensitivity of delta to underlying	$\frac{e^{-D(T-t)}N'(d_1)}{\sigma S\sqrt{T-t}}$
Theta $\frac{\partial V}{\partial t}$ Sensitivity to time	$-\frac{\sigma Se^{-D(T-t)}N'(d_1)}{2\sqrt{T-t}}+DSN(d_1)e^{-D(T-t)}$ $-rKe^{-r(T-t)}N(d_2)$
Speed $\frac{\partial^3 V}{\partial S^3}$ Sensitivity of gamma to underlying	$-\frac{e^{-D(T-t)}N'(d_1)}{\sigma^2S^2(T-t)}\times\left(d_1+\sigma\sqrt{T-t}\right)$
Charm $\frac{\partial^2 V}{\partial S\,\partial t}$ Sensitivity of delta to time	$De^{-D(T-t)}N(d_1)+e^{-D(T-t)}N'(d_1)$ $\times\left(\frac{d_2}{2(T-t)}-\frac{r-D}{\sigma\sqrt{T-t}}\right)$
Colour $\frac{\partial^3 V}{\partial S^2\,\partial t}$ Sensitivity of gamma to time	$\frac{e^{-D(T-t)}N'(d_1)}{\sigma S\sqrt{T-t}}$ $\times\left(D+\frac{1-d_1d_2}{2(T-t)}-\frac{d_1(r-D)}{\sigma\sqrt{T-t}}\right)$
Vega $\frac{\partial V}{\partial \sigma}$ Sensitivity to volatility	$S\sqrt{T-t}e^{-D(T-t)}N'(d_1)$
Rho (r) $\frac{\partial V}{\partial r}$ Sensitivity to interest rate	$K(T-t)e^{-r(T-t)}N(d_2)$
Rho (D) $\frac{\partial V}{\partial D}$ Sensitivity to dividend yield	$-(T-t)Se^{-D(T-t)}N(d_1)$
Vanna $\frac{\partial^2 V}{\partial S\,\partial \sigma}$ Sensitivity of delta to volatility	$-e^{-D(T-t)}N'(d_1)\frac{d_2}{\sigma}$
Volga/Vomma $\frac{\partial^2 V}{\partial \sigma^2}$ Sensitivity of vega to volatility	$S\sqrt{T-t}e^{-D(T-t)}N'(d_1)\frac{d_1d_2}{\sigma}$

Table 6.2: Formulæ for European put.

	Put
Payoff	$\max(K-S,0)$
Value V Black–Scholes value	$-Se^{-D(T-t)}N(-d_1)+Ke^{-r(T-t)}N(-d_2)$
Delta $\frac{\partial V}{\partial S}$ Sensitivity to underlying	$e^{-D(T-t)}(N(d_1)-1)$
Gamma $\frac{\partial^2 V}{\partial S^2}$ Sensitivity of delta to underlying	$\frac{e^{-D(T-t)}N'(d_1)}{\sigma S\sqrt{T-t}}$
Theta $\frac{\partial V}{\partial t}$ Sensitivity to time	$-\frac{\sigma Se^{-D(T-t)}N'(-d_1)}{2\sqrt{T-t}}-DSN(-d_1)e^{-D(T-t)}$ $+rKe^{-r(T-t)}N(-d_2)$
Speed $\frac{\partial^3 V}{\partial S^3}$ Sensitivity of gamma to underlying	$-\frac{e^{-D(T-t)}N'(d_1)}{\sigma^2S^2(T-t)}\times\left(d_1+\sigma\sqrt{T-t}\right)$
Charm $\frac{\partial^2 V}{\partial S\,\partial t}$ Sensitivity of delta to time	$De^{-D(T-t)}(N(d_1)-1)+e^{-D(T-t)}N'(d_1)$ $\times\left(\frac{d_2}{2(T-t)}-\frac{r-D}{\sigma\sqrt{T-t}}\right)$
Colour $\frac{\partial^3 V}{\partial S^2\,\partial t}$ Sensitivity of gamma to time	$\frac{e^{-D(T-t)}N'(d_1)}{\sigma S\sqrt{T-t}}$ $\times\left(D+\frac{1-d_1d_2}{2(T-t)}-\frac{d_1(r-D)}{\sigma\sqrt{T-t}}\right)$
Vega $\frac{\partial V}{\partial \sigma}$ Sensitivity to volatility	$S\sqrt{T-t}e^{-D(T-t)}N'(d_1)$
Rho (r) $\frac{\partial V}{\partial r}$ Sensitivity to interest rate	$-K(T-t)e^{-r(T-t)}N(-d_2)$
Rho (D) $\frac{\partial V}{\partial D}$ Sensitivity to dividend yield	$(T-t)Se^{-D(T-t)}N(-d_1)$
Vanna $\frac{\partial^2 V}{\partial S\,\partial \sigma}$ Sensitivity of delta to volatility	$-e^{-D(T-t)}N'(d_1)\frac{d_2}{\sigma}$
Volga/Vomma $\frac{\partial^2 V}{\partial \sigma^2}$ Sensitivity of vega to volatility	$S\sqrt{T-t}e^{-D(T-t)}N'(d_1)\frac{d_1d_2}{\sigma}$

'Hedging' in its broadest sense means the reduction of risk by exploiting relationships or correlation (or lack of correlation) between various risky investments. The purpose behind hedging is that it can lead to an improved risk/return.

Model-independent hedging: An example of such hedging is put-call parity. There is a simple relationship between calls and puts on an asset (when they are both European and with the same strikes and expiries), the underlying stock and a zero-coupon bond with the same maturity. This relationship is completely independent of how the underlying asset changes in value. Another example is spot-forward parity. In neither casedo we have to specify the dynamics of the asset, not even its volatility, to find a possible hedge. Such model-independent hedges are few and far between.

Model-dependent hedging: Most sophisticated finance hedging strategies depend on a model for the underlying asset. The obvious example is the hedging used in the Black–Scholes analysis that leads to a whole theory for the value of derivatives. In pricing derivatives we typically need to at least know the volatility of the underlying asset. If the model is wrong then the option value and any hedging strategy could also be wrong.

Delta hedging One of the building blocks of derivatives theory is delta hedging. This is the theoretically perfect elimination of all risk by using a very clever hedge between the option and its underlying. Delta hedging exploits the perfect correlation between the changes in the option value and the changes in the stock price. This is an example of 'dynamic' hedging; the hedge must be continually monitored and frequently adjusted by the sale or purchase of the underlying asset. Because of the frequent rehedging, any dynamic hedging strategy is going to result in losses due to transaction costs. In some markets this can be very important.

Gamma hedging To reduce the size of each rehedge and/or to increase the time between rehedges, and thus reduce costs, the technique of gamma hedging is often employed. A portfolio that is delta hedged is insensitive to movements in the underlying as long as those movements are quite small. There is a small error in this due to the convexity of the portfolio with respect to the underlying. Gamma hedging is a more accurate form of hedging that theoretically eliminates these second-order effects. Typically, one hedges one, exotic, say, contract with a vanilla contract and the underlying. The quantities of the vanilla and the underlying are chosen so as to make both the portfolio delta and the portfolio gamma instantaneously zero.

Vega hedging The prices and hedging strategies are only as good as the model for the underlying. The key parameter that determines the value of a contract is the volatility of the underlying asset. Unfortunately, this is a very difficult parameter to measure. Nor is it usually a constant as assumed in the simple theories. Obviously, the value of a contract depends on this parameter, and so to ensure that a portfolio value is insensitive to this parameter we can vega hedge. This means that we hedge one option with both the underlying and another option in such a way that both the delta and the vega, the sensitivity of the portfolio value to volatility, are zero. This is often quite satisfactory in practice but is usually theoretically inconsistent; we should not use a constant volatility (basic Black–Scholes) model to calculate sensitivities to parameters that are assumed not to vary. The distinction between variables (underlying asset price and time) and parameters (volatility, dividend yield, interest rate) is extremely important here. It is justifiable to rely on sensitivities of prices to variables, but usually not sensitivity to parameters. To get around this problem it is possible to independently model volatility, etc., as variables themselves. In such a way it is possible to build up a consistent theory.

Superhedging In incomplete markets you cannot eliminate all risk by classical dynamic delta hedging. But sometimes you can superhedge meaning that you construct a portfolio that has a positive payoff whatever happens to the market. A simple example of this would be to superhedge a short call position by buying one of the stock, and never rebalancing. Unfortunately, as you can probably imagine, and certainly as in this example, superhedging might give you prices that differ vastly from the market.

Crash (Platinum) hedging The final variety of hedging is specific to extreme markets. Market crashes have at least two obvious effects on our hedging. First of all, the moves are so large and rapid that they cannot be traditionally delta hedged. The convexity effect is not small. Second, normal market correlations become meaningless. Typically all correlations become one (or minus one). Crash or Platinum hedging exploits the latter effect in such a way as to minimize the worst possible outcome for the portfolio. The method, called CrashMetrics, does not rely on parameters such as volatilities and so is a very robust hedge. Platinum hedging comes in two types: hedging the paper value of the portfolio and hedging the margin calls.

4. Quick Overview of Mathematical Models for Finance

CIR Model, Heston Model, HJM etc.

5. Interest Rate Models

From one factor short rate to two factor short rate to Heath-Jarrow-Morton Framework.

Market Models: LFM, LSM

Consider a perfect market, i.e., all bonds with all maturities exist. There is only one interest rate for lending and for borrowing.

$P(t,T)$ denotes the price at t of the zero-coupon bond with maturity T. Of course, $P(T,T)=1$. Yield to maturity is denoted as $Y(t,T)$. $P(t,T)$ is $\mathcal{F}_t$-measurable and it is always positive and continuously differentiable.

A forward contract is an agreement to buy at T the asset S for $S_F(0,T)$. It is designed to enter into with no premium cost. The payoff for forward contract at T is $S_T - S_F(0,T)$. In order to replicate the payoff of the forward contract, a buy&hold strategy until T is to buy 1 share of stock at time 0 for S_0 and sell $S_F(0,T)$ units of bonds for $P(0,T)$. At time T, the payoff of this portfolio is $S_T - S_F(0,T)$ which is identical with the payoff of the forward contract. By law of one price principle, the price of the strategy is exactly the same to the price of the forward, i.e.,

$$S_0 - S_F(0,T)\,P(0,T) = 0$$

Solving above equation, we get the forward price (strike price)

$$S_F(0,T) = \frac{S_0}{P(0,T)}$$

and the forward price at time t is

$$S_F(t,T) = \frac{S_t}{P(t,T)}$$

Similarly, **a forward rate** for $[S,T]$with $0<S<T$ is a contract guaranteeing a risk-free rate. We want to replicate a contract that allows us to invest 1 at time S and get $R(0,S,T)$ interest rate for the period $[S,T]$. It is designed at 0 with a 0 premium.

There are two cash flows: at time S, we invest 1 dollar and at time T, we receive $e^{R(0,S,T)(T-S)}$in return.

To replicate the above payoff, we devise a strategy:

- At time 0, sell a bond with maturity S (receive $P(0,S)$instantly), buy bond with maturity T for $\#\frac{P(0,S)}{P(0,T)}$ shares.
- At time S, must pay 1 dollar
- At time T, you receive $\frac{P(0,S)}{P(0,T)}$ dollars.

Therefore let $e^{R(0,S,T)(T-S)} = \frac{P(0,S)}{P(0,T)}$, we can get

$$R(0,S,T) = \frac{1}{T-S}\left(\ln P(0,S) - \ln P(0,T)\right)$$

Similarly, a forward rate is given by

$$R(t,S,T) = \frac{-\left(\ln P(t,T) - \ln P(t,S)\right)}{T-S}$$

and when $T\to S$, we get the **instantaneous forward rate**

$$R(t,S,T) \to -\frac{\partial \ln P(t,T)}{\partial T} = \mathcal{F}(t,T)$$

If $\mathcal{F}(t,T)$ is known for all t and T,

$$P(t,T) = \exp\left(-\int_t^T \mathcal{F}(t,u)\,du\right)$$

Forward rates characterize the bond prices. Yield to maturity $P(t,T) = \exp\left(-Y(t,T)(T-t)\right)$thus

$$Y(t,T) = \frac{\int_t^T \mathcal{F}(t,u)\,du}{T-t} = -\frac{\ln P(t,T)}{T-t}$$

$r(t)$ is the yield of investment at time t for an infinitesimal time period and

$$Y(t; t+dt) = -\frac{\ln P(t, t+dt)}{dt} = \frac{\partial}{\partial T}(-\ln P(t, T))|_{T=t} = \mathcal{F}(t, t)$$

Note that

$$B(t) = \exp\left(\int_0^t r_s ds\right)$$

ie, bank account with stochastic interest rates. If we assume there is no arbitrage in the market, there are strong relationships among $dP(t,T), d\mathcal{F}(t,T), dr_t, dB_t$

Linkage between $P(t,T)$ and r_t:

If r_t is constant, $P(t,T) = \exp(-r(T-t))$ .If not, is it true $P(t,T) = \exp\left(-\int_t^T r_s ds\right)$? If there is no arbitrage, there must exist risk neutral measure $\mathbb{Q}$ such that

$$P(t,T) = \mathbb{E}_{\mathbb{Q}}\left[\exp\left(-\int_t^T r_s ds\right)|\mathcal{F}_t\right]$$

Term Structure of Interest Rate

Short Rates Model:

$$dr_t = \mu(t, r_t)\, dt + \sigma(t, r_t)\, dW_t$$

notice that r is not a traded asset. The bank account

$$dB_t = r_t B_t dt$$

and price of ZC bond

$$P(t,T) = F(t, r_t, T) := F^T(t, r)$$

can not be replicated with B because there is no unique price for P. There is a strong relationship between $P(t,T)$ and $P(t,S)$. The idea to find the link is to build a portfolio with $P(t,T)$ and $P(t,S)$ to replicate a bank account. It is equivalent to building an adapted self-financing strategy that is risk free.

$$V(t) = h_1(t) P(t,T) + h_2(t) P(t,S)$$

where $h_1(t)$ denotes the number of shares of bond with maturity T and $h_2(t)$ denotes the number of shares of bond with maturity S. We want to make V riskless and self-financing.

$$dV(t) = h_1 dP(t,T) + h_2 dP(t,S)$$

By Itô's formula,

$$\begin{aligned} dP(t,T) &= F_t^T dt + F_r^T dr + \frac{1}{2} F_{rr}^T (dr)^2 \\ &= F_t^T dt + +F_r^T(\mu dt + \sigma dW) + \frac{1}{2} F_{rr}^T \sigma^2 dt \\ &= \left(F_t^T + \mu F_r^T + \frac{1}{2}\sigma^2 F_{rr}^T\right) dt + \sigma F_r^T dW \end{aligned}$$

hence

$$\frac{dP(t,T)}{P(t,T)} = \alpha^T dt + \sigma^T dW$$

where $\alpha^T := \frac{F_t^T + \mu F_r^T + \frac{1}{2}\sigma^2 F_{rr}^T}{F^T}$ and $\sigma^T := \frac{\sigma F_r^T}{F^T}$. Thus $\frac{dV(t,T)}{V(t)}$ can be represented as

$$\begin{aligned}\frac{dV(t,T)}{V(t)} &= \frac{h_1 P(t,T)\left(\alpha^T dt + \sigma^T dW\right) + h_2 P(t,S)\left(\alpha^S dt + \sigma^S dW\right)}{V(t)} \\ &= \left(\underbrace{\frac{h_1 P(t,T)}{V(t)}}_{u_t}\alpha^T + \underbrace{\frac{h_2 P(t,S)}{V(t)}}_{v_t}\alpha^S\right) dt \\ &\quad + \left(\frac{h_1 P(t,T)}{V(t)}\sigma^T + \frac{h_2 P(t,S)}{V(t)}\sigma^S\right) dW\end{aligned}$$

Notice that $u_t + v_t = 1$ and we want $V(t)$ riskless, i.e., $u_t\sigma^T + v_t\sigma^S = 0$ therefore

$$u_t = -\frac{\sigma^S}{\sigma^T - \sigma^S}, v_t = \frac{\sigma^T}{\sigma^T - \sigma^S}$$

The no arbitrage condition implies that

$$u_t\alpha^T + v_t\alpha^S = r_t$$

which is equivalent to

$$\frac{-\sigma^S}{\sigma^T - \sigma^S}\alpha^T + \frac{\sigma^T}{\sigma^T - \sigma^S}\alpha^S = r_t$$

$\forall S, T$,

$$\frac{\alpha^T - r_t}{\sigma^T} = \frac{\alpha^S - r_t}{\sigma^S} = \lambda(t)$$

this ratio is the **market price of interest rate risk** and it is independent of T and S. In conclusion, we arrive at the PDE for bond price

$$F_t^T + (\mu - \lambda\sigma) F_r^T + \frac{1}{2}\sigma^2 F_{rr}^T - rF^T = 0$$

with the boundary condition $F^T(T,R) = 1$ where $\alpha^T = \frac{F_t^T + \mu F_r^T + \frac{1}{2}F_{rr}^T\sigma^2}{F^T}$, $\sigma^T = \frac{\sigma F_r^T}{F^T}$ and $\lambda = \frac{\alpha^T - r}{\sigma^T}$.

Moreover,

$$F^T(t, r_t, T) = \mathbb{E}_{\mathbb{Q}}\left[e^{-\int_t^T r_s ds}\right]$$

where

$$dr_t = (\mu - \lambda\sigma) dt + \sigma dZ_t$$

where Z is a BM under $\mathbb{Q}$ and each λ gives a risk neutral measure.

The **Vasicek model** (one factor model describing the evolution of interest rate) is given by

$$dr_t = \underbrace{(b - ar)}_{\mu(t,r)} dt + \underbrace{\sigma}_{\sigma(t,r)} dW$$

By Itô's lemma, we can get

$$\begin{aligned}d\left(r_t e^{at}\right) &= are^{at}dt + e^{at}\left((b - ar)dt + \sigma dW\right) + 0 \\ &= be^{at}dt + \sigma e^{at}dW\end{aligned}$$

Taking integration on both sides, we get

$$r_T e^{aT} - r_0 = \int_0^T be^{at}dt + \sigma\int_0^T e^{at}dW_t$$

thus

$$r_T = r_0 e^{-at} + e^{-aT}\left(\int_0^T b e^{at} dt + \underbrace{\sigma e^{-aT}\int_0^T e^{at} dW}_{\sim N\left(0,\sigma^2 e^{-2aT}\int_0^T e^{2at}dt\right)}\right)$$

CIR model (Cox-Ingersoll-Ross)

$$dr = a\left(b-r\right)dt + \sigma\sqrt{r}dW$$

It can be proved that $r > 0$ almost surely.

Hull-White Model

$$dr = \left(\theta(t) - ar\right)dt + \sigma dW$$

Dothan Model

$$dr = ardt + \sigma r dW$$

hence

$$r_t = r_0 \exp\left(\left(a - \frac{\sigma^2}{2}\right)t + \sigma W_t\right) \geq 0$$

In a one factor model, $cor\left(P\left(t,T_1\right), P\left(t,T_2\right)\right) = 1$ which is not true in a two-factor model.

HJM model Refer to Wikipedia

The classical short rate models are shown in the following table: From "Interest Rate Models - Theory and Practice" by Brigo

Model	Dynamics	$r>0$	$r\sim$	AB	AO
V	$dr_t = k[\theta - r_t]dt + \sigma dW_t$	N	$\mathcal{N}$	Y	Y
CIR	$dr_t = k[\theta - r_t]dt + \sigma\sqrt{r_t}dW_t$	Y	NCχ^2	Y	Y
D	$dr_t = ar_t dt + \sigma r_t dW_t$	Y	L$\mathcal{N}$	Y	N
EV	$dr_t = r_t\left[\eta - a\ln r_t\right]dt + \sigma r_t dW_t$	Y	L$\mathcal{N}$	N	N
HW	$dr_t = k[\theta_t - r_t]dt + \sigma dW_t$	N	$\mathcal{N}$	Y	Y
BK	$dr_t = r_t\left[\eta_t - a\ln r_t\right]dt + \sigma r_t dW_t$	Y	L$\mathcal{N}$	N	N
MM	$dr_t = r_t\left[\eta_t - \left(\lambda - \frac{\gamma}{1+\gamma t}\right)\ln r_t\right]dt + \sigma r_t dW_t$	Y	L$\mathcal{N}$	N	N
CIR++	$r_t = x_t + \varphi_t,\ \ dx_t = k[\theta - x_t]dt + \sigma\sqrt{x_t}dW_t$	Y*	SNCχ^2	Y	Y
EEV	$r_t = x_t + \varphi_t,\ dx_t = x_t[\eta - a\ln x_t]dt + \sigma x_t dW_t$	Y*	SL$\mathcal{N}$	N	N

Table 3.1. Summary of instantaneous short rate models.

Where V, CIR, D, EV, HW, BK, MM, CIR++, EVV stand respectively for the Vasicek (1977) model, the Cox, Ingersoll and Ross (1985) model, the Dothan (1978) model, the Exponential Vasicek model, the Hull and White (1990) model, the Black and Karasinski (1991) model, the Mercurio and Moraleda (2000) model, the CIR++ model and the Extended Exponential Vasicek model. N and Y stand respectively for "No" and "Yes", whereas Y$*$ means that rates are positive under suitable conditions for the deterministic function φ; $\mathcal{N}$, $\mathcal{LN}$, $\mathcal{NC}\chi^2$, $\mathcal{SNC}\chi^2$, $\mathcal{SLN}$ denote respectively normal, lognormal, noncentral χ^2, shifted noncentral χ^2 and shifted lognormal distributions; AB(O) stands for Analytical Bond (Option) price.

Interest rate swap is a contract between two parties to exchange interest on a specified principal. The exchange may be fixed for floating or floating of one tenor for floating of another tenor. Fixed for floating is a particularly common form of swap. These instruments are used to convert a fixed-rate loan to floating, or vice versa. Usually the interval between the exchanges is set to be the same as the tenor of the floating leg. Furthermore, the floating leg is set at the payment date before it is paid. This means that each floating leg is equivalent to a deposit and a withdrawal of the principal with an interval of the tenor between them. Therefore all the floating legs can be summed up to give one deposit at the start of the swap's life and a withdrawal at maturity. This means that swaps can be valued directly from the yield curve without needing a dynamic model. When the contract is first entered into the fixed leg is set so that the swap has zero value. The fixed leg of the swap is then called the par swap rate and is a commonly quoted rate. These contracts are so liquid that they define the longer-maturity end of the yield curve rather than vice versa.

Mortgage Backed Security (MBS) is a pool of mortgages that have been securitized. All of the cashflows are passed on to investors, unlike in the more complex CMOs. The risks inherent in MBSs are interest rate risk and prepayment risk, since the holders of mortgages have the right to prepay. Because of this risk the yield on MBSs should be higher than yields without prepayment risk. Prepayment risk is usually modelled statistically, perhaps with some interest rate effect. Holders of mortgages have all kinds of reasons for prepaying, some rational and easy to model, some irrational and harder to model but which can nevertheless be interpreted statistically.

Straddle is a portfolio consisting of a long call and a long put with the same strike and expiration. Such a portfolio is for taking a view on the range of the underlying or volatility.

Swaption is an option on a swap. It is the option to enter into the swap at some expiration date, the swap having predefined characteristics. Such contracts are very common in the fixed income world where a typical swaption would be on a swap of fixed for floating. The contract may be European so that the swap can only be entered into on a certain date, or American in which the swap can be entered into before a certain date or Bermudan in which there are specified dates on which the option can be exercised.

6. Several Pricing Examples

6.1. **Call on Call.** Call option will maturity T_1, strike L and the underlying is a call option with strike K and a maturity $T_2 (T_1 < T_2)$. The payoff of this derivative is given by$\left(C\left(T_1, S_{T_1}, K, T_2\right) - L\right)^+$and this call on call option can be priced using risk-neutral method. The price at time 0 is given by

$$\begin{aligned} & \mathbb{E}_{\mathbb{Q}}\left[e^{-rT_1}\left(C\left(T_1, S_{T_1}, K, T_2\right) - L\right)^+\right] \\ = \quad & e^{-rT_1}\mathbb{E}_{\mathbb{Q}}\left[C\left(T_1, S_{T_1}, K, T_2\right)\cdot \mathbf{1}_{\left\{C\left(T_1, S_{T_1}, K, T_2\right)>L\right\}} - L\cdot \mathbf{1}_{\left\{C\left(T_1, S_{T_1}, K, T_2\right)>L\right\}}\right] \end{aligned}$$

Notice that

$$\begin{aligned} C\left(T_1, S_{T_1}, K, T_2\right) &= e^{-r(T_2-T_1)}\mathbb{E}_{\mathbb{Q}}\left[\left(S_{T_2} - K\right)^+ | \mathcal{F}_{T_1}\right] \\ &= S_{T_1}\mathcal{N}\left(d_1\right) - ke^{-r(T_2-T_1)}\mathcal{N}\left(d_2\right) \end{aligned}$$

where

$$d_1 = \frac{\ln\left(\frac{S_{T_1}}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)(T_2 - T_1)}{\sigma\sqrt{T_2 - T_1}}$$

and

$$d_2 = d_1 - \sigma\sqrt{T_2 - T_1}$$

Keep in mind that $C(T_1, S_{T_1}, K, T_2)$ is a random variable at time 0 but it is $\mathcal{F}_{T_1}$-measurable. Let $f(S_{T_1})$ denote this option value. (f is strictly increasing)

Hence

$$\{C(T_1, S_{T_1}, K, T_2) > L\} = \left\{S_{T_1} > f^{-1}(L)\right\}$$

If $T_2 = T_1$, $\left\{(S_{T_2} - K)^+ > L\right\} = \{S_{T_2} > K + L\}$ and

$$\frac{\partial C(T_1, S_{T_1}, K, T_2)}{\partial S_{T_1}} = \mathcal{N}(d_1) \in (0,1) > 0$$

by observing that (Delta of a call option)

$$\frac{\partial C}{\partial S_{T_1}} = \mathcal{N}(d_1) + \underbrace{S_{T_1}\mathcal{N}'(d_1) \cdot \frac{\partial d_1}{\partial S_{T_1}} - ke^{-r(T_2-T_1)}\mathcal{N}'(d_2)\frac{\partial d_2}{\partial S_{T_1}}}_{=0}$$

where $\mathcal{N}'(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$. Let $S^\star$ be such that $C(T_1, S^*, K, T_2) = L$ which can be found numerically. Then the price of the call on call is

$$e^{-rT_1}\mathbb{E}_{\mathbb{Q}}\left[C(T_1)\,\mathbf{1}_{\left\{S_{T_1} > S^\star\right\}}\right] - e^{-rT_1}L \cdot \mathbb{Q}(S_{T_1} > S^\star)$$

Note that

$$\begin{aligned}
\mathbb{Q}(S_{T_1} > S^\star) &= \mathbb{Q}\left(S_0 e^{\left(r - \frac{\sigma^2}{2}\right)T_1 + \sigma T_1} > S^\star\right) \\
&= \mathbb{Q}\left(\frac{W_{T_1}}{\sqrt{T_1}} > \frac{\ln(S^\star/S_0) - \left(r - \frac{\sigma^2}{2}\right)T_1}{\sigma\sqrt{T_1}}\right) \\
&= \mathcal{N}\left(\frac{\ln(S_0/S^\star) + \left(r - \frac{\sigma^2}{2}\right)T_1}{\sigma\sqrt{T_1}}\right)
\end{aligned}$$

Recall that $\mathbb{E}[\mathbb{E}[Y|\mathcal{F}]] = \mathbb{E}[Y]$ hence as a direct result, if X is $\mathcal{F}$-measurable, $\mathbb{E}[X\mathbb{E}[Y|\mathcal{F}]] = \mathbb{E}[XY]$. In regard to the first term

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}\left[C(T_1) \cdot \mathbf{1}_{\left\{S_{T_1} > S^\star\right\}}\right] &= \mathbb{E}_{\mathbb{Q}}\left[e^{-r(T_2-T_1)} \cdot \mathbf{1}_{\left\{S_{T_1} > S^\star\right\}} \cdot \mathbb{E}_{\mathbb{Q}}\left[(S_{T_2} - K)^+ |\mathcal{F}_{T_1}\right]\right] \\
&= e^{-r(T_2-T_1)}\mathbb{E}\left[\mathbf{1}_{\left\{S_{T_1} > S^\star\right\}} \cdot (S_{T_2} - K)^+\right] \\
&= e^{-r(T_2-T_1)}\underbrace{\mathbb{E}\left[\mathbf{1}_{\left\{S_{T_1} > S^\star\right\}} \cdot S_{T_2} \cdot \mathbf{1}_{\left\{S_{T_2} > K\right\}}\right]}_{A} - e^{-r(T_2-T_1)}\underbrace{\mathbb{Q}(S_{T_1} > S^\star, S_{T_2} > K)}_{B}
\end{aligned}$$

where

$$B = \mathbb{Q}\left(\ln S_0 + \left(r - \frac{\sigma^2}{2}\right)T_1 + \sigma W_{T_1} > \ln S^\star, \ln S_0 + \left(r - \frac{\sigma^2}{2}\right)T_2 + \sigma W_{T_2} > \ln K\right)$$

Define

$$\mathcal{N}(a,b,\rho)=\mathbb{P}(X\le a,Y\le b)$$

where $X\sim\mathcal{N}(0,1), Y\sim\mathcal{N}(0,1)$ with correlation factor $\rho=cov(X,Y)$. Thus

$$B=\mathbb{Q}\left(\frac{W_{T_1}}{\sqrt{T_1}}>\frac{\ln(S^\star/S_0)-\left(r-\frac{\sigma^2}{2}\right)T_1}{\sigma\sqrt{T_1}},\frac{W_{T_1}}{\sqrt{T_1}}>\frac{\ln(K/S_0)-\left(r-\frac{\sigma^2}{2}\right)T_2}{\sigma\sqrt{T_2}}\right)$$

and notice that

$$\begin{aligned}\mathbb{Q}(X>c,Y>d) &= 1-\mathbb{Q}(X\le c,Y<\infty)-\mathbb{Q}(X<\infty,Y\le d)+\mathbb{Q}(X\le c,Y\le d)\\ &= 1-\mathcal{N}(c,\infty,\rho)-\mathcal{N}(\infty,d,\rho)+\mathcal{N}(c,d,\rho)\end{aligned}$$

where $c=\frac{\ln(S^\star/S_0)-\left(r-\frac{\sigma^2}{2}\right)T_1}{\sigma\sqrt{T_1}}$ and $d=\frac{\ln(K/S_0)-\left(r-\frac{\sigma^2}{2}\right)T_2}{\sigma\sqrt{T_2}}$.

$$\begin{aligned}A: &= \mathbb{E}\left[\mathbf{1}_{\left\{S_{T_1}>S^\star\right\}}\cdot S_{T_2}\cdot\mathbf{1}_{\left\{S_{T_2}>K\right\}}\right]\\ &= \mathbb{E}\left[S_0e^{\left(r-\frac{\sigma^2}{2}\right)T_2+\sigma T_2}\mathbf{1}_{\left\{S_{T_1}>S^\star\right\}}\cdot\mathbf{1}_{\left\{S_{T_2}>K\right\}}\right]\\ &= S_0e^{rT_2}\mathbb{E}\left[\frac{d\tilde{\mathbb{Q}}}{d\mathbb{Q}}\cdot 1_{\left\{S_{T_1}>S^\star,S_{T_2}>K\right\}}\right]\end{aligned}$$

By Girsanov theorem, $\tilde{W}$ is a BM under $\tilde{\mathbb{Q}}$. ($\mathbb{E}_{\mathbb{Q}}\left[\frac{d\tilde{\mathbb{Q}}}{d\mathbb{Q}}X\right]=\mathbb{E}_{\tilde{\mathbb{Q}}}[X]$, which can be seen from $\mathbb{E}_{\mathbb{Q}}\left[\frac{d\tilde{\mathbb{Q}}}{d\mathbb{Q}}X\right]=\int\frac{d\tilde{\mathbb{Q}}}{d\mathbb{Q}}Xd\mathbb{Q}=\int Xdd\tilde{\mathbb{Q}}$, not rigorous). Thus

$$\begin{aligned}A &= S_0e^{rT_2}\mathbb{E}_{\tilde{\mathbb{Q}}}\left[1_{\left\{S_{T_1}>S^\star,S_{T_2}>K\right\}}\right]\\ &= S_0e^{rT_2}\tilde{\mathbb{Q}}\left(\frac{\tilde{W}_{T_1}}{\sqrt{T_1}}>\frac{\ln(S^\star/S_0)-\left(r-\frac{\sigma^2}{2}\right)T_1}{\sigma\sqrt{T_1}},\frac{\tilde{W}_{T_1}}{\sqrt{T_1}}>\frac{\ln(K/S_0)-\left(r-\frac{\sigma^2}{2}\right)T_2}{\sigma\sqrt{T_2}}\right)\end{aligned}$$

6.2. **Exchange Option.** S_1 and S_2 are two assets. Exchange option gives the right to receive S_1 by paying S_2 at T. Thus the payoff function is given by

$$X_T=(S_1(T)-S_2(T))^+$$

where

$$\begin{cases}\frac{dS_1}{S_1}=rdt+\sigma_{11}dW_1+\sigma_{12}dW_2\\ \frac{dS_2}{S_2}=rdt+\sigma_{21}dW_1+\sigma_{22}dW_2\end{cases}$$

The price of the exchange option can be derived by computing

$$\mathbb{E}\left[e^{-rT}(S_1(T)-S_2(T))^+\right]$$

or make use of the following theorem since $\Phi\left(\overrightarrow{S}\right)=\max(S_1-S_2,0)$ is homogeneous of degree 1.

Theorem 3. *(Theorem of Reduction) Let $\overrightarrow{S}=(S_1,\cdots,S_n)$ and $\Phi\left(\lambda\overrightarrow{S}\right)=\lambda\Phi\left(\overrightarrow{S}\right)$. Then the price of the derivative at t is given by*

$$F\left(t,\overrightarrow{S_t}\right)=S_nG\left(t,\frac{S_1}{S_n},\cdots,\frac{S_{n-1}}{S_n}\right)$$

where G solves

$$\begin{aligned} 0 &= G_t + \frac{1}{2}\sum_{i=1}^{n-1}\sum_{j=1}^{n-1}\xi_i\xi_j G_{ij}\left(t,\xi_1,\cdots,\xi_{n-1}\right)D_{ij} \\ G\left(t,\overrightarrow{\xi}\right) &= \Phi\left(\xi_1,\cdots,\xi_{n-1},1\right) \\ G_{ij} &= \frac{\partial^2 G}{\partial\xi_i\partial\xi_j} \\ D_{ij} &= C_{ij}+C_{nn}+C_{in}+C_{nj} \end{aligned}$$

and

$$C = \Sigma\Sigma^T$$

where

$$\Sigma = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1n} \\ \vdots & \ddots & \vdots \\ \sigma_{n1} & \cdots & \sigma_{nn} \end{pmatrix}$$

and

$$\frac{dS_i}{S_i} = \mu_i dt + \sum_j \sigma_{ij}dW_j$$

For exchange option, the terminal payoff is $F\left(T,S_1,S_2\right)=\max\left(S_1-S_2,0\right)$. In general, the price of the exchange option at time t is given by

$$F\left(t,S_1\left(t\right),S_2\left(t\right)\right) = S_2\left(t\right)G\left(t,\frac{S_1\left(t\right)}{S_2\left(t\right)}\right)$$

and

$$G_t + \frac{1}{2}\xi^2 G_{\xi\xi}\left(t,\xi\right)D = 0$$

and

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 \\ \sigma_2\rho & \sigma_2\sqrt{1-\rho^2} \end{pmatrix}$$

thus

$$\Sigma\Sigma^T = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

Thus

$$D = \sigma_1^2+\sigma_2^2+2\rho\sigma_1\sigma_2$$

Therefore we get

$$\begin{aligned} G_t + \frac{1}{2}\xi^2 G_{\xi\xi}D &= 0 \\ G\left(t,\xi\right) = F\left(t,\xi,1\right) &= \left(\xi-1\right)^+ \end{aligned}$$

Hence

$$G_0 = \mathbb{E}_{\mathbb{Q}}\left[e^{-rT}\left(S_T-1\right)^+\right]$$

and

$$\frac{dS_T}{S_T} = rdt + \sqrt{D}dW$$

Note that the standard Black-Scholes PDE is given by

$$-rC + C_t + rSC_S + \frac{1}{2}\sigma^2 S^2 C_{SS} = 0$$

hence $r = 0, K = 1, \sigma = \sqrt{D}$ in

$$G_t + \frac{1}{2}\xi^2 G_{\xi\xi} D = 0$$

In conclusion,

$$G(0, \xi_0) = \xi_0 \mathcal{N}(d_1) - \mathcal{N}(d_2)$$

where $d_1 = \frac{\ln \xi_0 + \frac{DT}{2}}{\sqrt{DT}}$ and $d_2 = d_1 - \sigma\sqrt{T}$.

Finally,

$$F_0 = S_2(0)\, G_0\left(0, \frac{S_1(0)}{S_2(0)}\right) = S_1(0)\,\mathcal{N}(d_1) - S_2(0)\,\mathcal{N}(d_2)$$

where $d_1 = \frac{\ln(S_1(0)/S_2(0)) + \frac{DT}{2}}{\sqrt{DT}}$ and $d_2 = d_1 - \sigma\sqrt{T}$.

Bear in mind that $d\overrightarrow{S} = r\overrightarrow{S}dt + \overrightarrow{S}\Sigma d\overrightarrow{W}$

6.3. **Change of Numéraire.** See reference books on the list.