SMM272 Risk Analysis Revision Notes

Heqing Shi

MSc Financial Mathematics 2019/2020

Contents

0 What to Know for the Exam

- to use transformation of variable: eg. given X and its cdf/pdf how to find the cdf/pdf of Y=g(X)
- The model r=m+sg*z and its properties (eg. mean, variance, simulation of cumulative returns)
- The Variance Ratio Test
- The model r= rho*r(t-1)+sg*z and its properties (eg. mean, variance, simulation of cumulative returns). How to adjust VaR for autocorrelation
- Parametric VaR: asset and portfolio level (covariance matrix); Cholesky decomposition and Monte Carlo simulation
- Non Parametric VaR: quantile, bootstrap/historical simulation
- Estimation Risk: Parametric vs Bootstrap
- Backtesting VaR: Kupiec test, its functioning and its limits. How to compute the probability of type 1 and type 2 errors
- Risk decomposition: MVaR, CVaR, IVaR
- Coherent Risk measures
- EWMA model: univariate and multivariate specfication. Pros & Cons. Comparison with GARCH and SV models.
- VaR & Derivatives: exact formula, delta var and delta-gamma
- Principal Component Analysis (PCA is also part of Fixed Income): derivation of the characteristic equation, computation of PCA in Excel and Matlab, how to use PCA to estimate the VaR of a bond portfolio
- Content of the coursework
- Exercise Handbook: Questions with up to 3 stars

- Exam Structure:

  Remember that simple calculations do not deserve full marks. You have to outline all the assumptions/relevant aspects that allow you to arrive to the final result.

Last modified: Monday, 13 April 2020, 10:37 AM

1 Introduction to Risk Analysis

Types of risks:

- Market Risk
- Credit Risk
- Model Risk
- Liquidity Risk
- Operational Risk

2 Measuring Return

- Model directly price, but it is non-stationary
- Model returns rather, as its is stationary
- The normality aasumption must be attached to log-returns rather to simple returns
- R (simple returns) will be distributed according to a shifted log-normal distribution
- Basic model for $\Delta-$period log-returns: $r(t) = r(t, t+\Delta) = \mu_\delta(t) + \sigma_\Delta(t)\epsilon(t)$
  - If stationary assumptions on r is valid, i.e. $\mathbb{C}ov_t(r(s, s+\Delta), r(u, u+\Delta)) = 0, \quad \forall u, s$
  - then $\mathbb{E}_t(r(t, t+\Delta)) = \mu_\Delta(t)$ and $\mathbb{V}_t(r(t, t+\Delta)) = \sigma^2_\Delta(t)$
- The assumption of zero-autocorrelation in log-returns can be tested using the Variacne Ratio Test
  - $\hat{\mathrm{VR}}(\mathrm{n}) = \frac{\hat{\sigma}^2_{\mathrm{cum}}}{\mathrm{n}\hat{\sigma}^2_\Delta} \sim \mathcal{N}(1, \frac{2(n-1)}{Tn})$, which is close to 1 if the assumption is true
  - normalize it, we have $z(n) = \frac{\hat{\mathrm{VR}}(n)-1}{\sqrt{\frac{2(n-1)}{Tn}}} \sim \mathcal{N}(0, 1)$
  - $[-1.96, 1.96]$ is the 95% acceptance region of the assumption
- MC simulation of returns: $r_\Delta(t) = \mu_\Delta + \sigma_\Delta\epsilon(t)$
- then stock prices: $S(t+\Delta) = S(t) \times e^{r_\Delta(t)}$
- Serially correlated AR(1) model:
  - $r(t) = \rho r(t-\Delta) + \sigma_\Delta\epsilon(t)$
  - standard deviation growths according to:

$$\begin{aligned}\mathbb{SD}ev(r(t, t+n\Delta)) &= \sigma_\Delta\sqrt{\frac{n(1-\rho^2) + \rho(1-\rho^n)(\rho^{n+1} - \rho - 2)}{(1-\rho)^3(1+\rho)}} \\ &= \sigma_\Delta\sqrt{\frac{n}{(1-\rho)^2}} \\ &= \sigma_\Delta\sqrt{n}\sqrt{\frac{1}{(1-\rho)^2}} \quad \text{(Approximation)}\end{aligned} \tag{1}$$

  - if we adopted square-root when the returns are AR(1), the ratio between true value and our estimate:

$$\frac{\sigma_\Delta\frac{\sqrt{n}}{1-\rho}}{\sigma_\Delta\sqrt{n}} = \frac{1}{1-\rho} \tag{2}$$

  - then we under-estimated the true volatility ($\rho > 0$) or over-estimated it ($\rho < 0$)
  - MC simulation of returns:

$$r_\Delta(t) = \rho r_\Delta(t-\Delta) + \sigma_\Delta\epsilon(t) \tag{3}$$

- Normal approximation is good for short horizons

3 Introduction to Value at Risk

- SPAN system not appropriate for muilt-assets portfolio
- If estimated properly, the P&L distribution reflects the netting and diversification effects and can be compared acorss portfolios
  - are past data useful to forecast the future risk?
  - best approach is a combination of approaches trying to include forward-looking information as well
- Basel committee chose:
  - a 10-day horizon
  - a 99% confidence level
  - the probability distribution construction is left to the user
  - Internal risk models: report 99% VaR over both one-day and 10-day horizon
  - $\mathrm{MRC(t) = VaR_{0.99}(t, t + \frac{10}{255}), (\frac{S(t)}{60} \sum_{i=0}^{59} VaR_{0.99}(t - \frac{i}{255}, t - \frac{i-10}{255}))}$

4 Estimating VaR: Parametric and Non-Parametric Approaches

- Parametric approach: such as the log-returns are distributed according to a Gaussian r.v., which requires parameters estimation
- Non-parametric approach: do not make any assumption and use past data to build the empirical distribution
- Gaussian parametric approach:
  - $ES_\alpha(t, t+\Delta) = -(\mu_\Delta - \frac{\sigma_\Delta}{1-\alpha}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(\frac{-VaR_\alpha - \mu_\Delta}{\sigma_\Delta})^2}) = -(\mu_\Delta - \sigma_\Delta \frac{\phi(Z_1 - \alpha)}{1-\alpha})$
  - $ES_\alpha(t, t+n\Delta) = -(\mu_\Delta - \sigma_\Delta \frac{\sigma_\Delta \sqrt{n}}{1-\alpha}\phi(Z_1 - \alpha))$
  - CONS:
    - assumes the future will be like the past
    - model-dependent procedure may be misleading if model is poor
    - cannot capture skewed and fat-tailed characteristic
- Historical Simulation Approach:
  - if $(1-\alpha)T$ is not an integer number, linear interpolation is required
  - no simple procedure to extrapolate a n-period VaR
  - need to collect n-period returns (inefficient)
  - BOOTSTRAPPING: for T sufficiently large, we expect the bootstrapped distribution to be near the true distribution (Central Limit Theorem)
  - not accurate compared with parametric Gaussian approach, but no exposure to model risk
  - sensitive to the length of data sample used
  - unconditional distribution, i.e. do not model changes in volatility
  - ghost effect
- Top-down approach:
  - specifies the portfolio P&L distribution without reference to the constituents
  - more parsimonious
  - not fine enough to identify the identity of the component that contributed the most to the portfolio VaR
- Bottom-up approach:
  - this requires the specification of the dependence structure, i.e. the joint distribution of the components

- Estimation error:
  - for large T, $s.e.(V\hat{a}R_{\alpha}(t, t+n\Delta)) = \sigma_{\Delta} \times \sqrt{\frac{n}{2T}} \times |Z_{1-\alpha}|$ (Parametric Approach)
  - for large T, $s.e.(V\hat{a}R_{1-\alpha}) = \sqrt{\frac{\alpha(1-\alpha)}{T\hat{f}^2(-V\hat{a}R_{1-\alpha})}}$ (Non-parametric Approach)
- log-normal distribution:
  - $\mathbb{E}(Y) = e^{\mu+\frac{1}{2}\sigma^2}$, and $\mathbb{V}ar(Y) = e^{2\mu+\sigma^2}(e^{\sigma^2-1})$

5 Regulation and Backtesting VaR

- Internal risk model:
  - report both the 99% VaR over a both a one-day horizon, i.e. $VaR_{0.99}(t, t+\frac{1}{255})$ and
  - the one on a 10-day horizon, i.e. $VaR_{0.99}(t, t+\frac{10}{255})$
- $\mathrm{MRC(t) = VaR_{0.99}(t, t+\frac{10}{255}), (\frac{S(t)}{60}\sum_{i=0}^{59} VaR_{0.99}(t-\frac{i}{255}, t-\frac{i-10}{255}))}$
- even when the multilier takes the value of 1.5, the MRC produces a much higher risk measure than 99.9% VaR
- Unconditional Coverage
  - The Kupiec Test:
    * examines the number of violations but does not consider if they cluster in time or not
    * also called POF test
    * $1-\alpha$ is the theoretical probability of having a violation
    * $1-\hat{\alpha} = \frac{\sum_{i=1}^{n} I_{i\Delta}(\alpha)}{n}$ is the observed violation frequency
    * log-likelihood ratio: $LR_{uc} = -2\log\frac{L(j,n,\alpha)}{L(j,n,\hat{\alpha})}$
    * null hypothesis: the VaR model is good (LR_{uc} should have values near to 0)
    * Asymptotically:

$$\begin{aligned} LR_{uc} &= -2(\log((\frac{1-\alpha}{1-\hat{\alpha}})^j) + \log((\frac{\alpha}{\hat{\alpha}})^{n-j})) \\ &= -2(j\log(\frac{1-\alpha}{1-\hat{\alpha}}) + (n-j)\log(\frac{\alpha}{\hat{\alpha}})) \quad \sim \mathcal{X}_1^2 \end{aligned} \tag{4}$$

    * Limits of Kupiec test:
      · require a large number of data
      · low power of test (Type II error) (increase the number of observations can reduce this problem)
      · low power problem is exhacerbated when using MC simulation as the true model is changing conditional volatility
      · focus only on the number of exceptions, without considering the time distribution of those exceptions
      · do not care about the size of the violation as well
- Independence property
  - previous VaR violations must not convey any information about whether or not an additional VaR violation
  - if a VaR violation is more likely to occur after a previous VaR violation, the this implies that the probability of $I_{t+\Delta}(\alpha)$ conditional on the event that $I_{\alpha}(a) = 1$ exceeds $1-\alpha$, and indicating that VaR estimate is too small

6 Coherent Risk Measure

- VaR does not describe the maximum loss
- VaR does not describe the losses in the left tail
- estimation risk in VaR (the sampling variability due to limited sample size)
  - larger sample, better accuracy
  - larger confidence level, lower accuracy
- may be not sub-additive
  - diversification benefits (sub-additivity): $VaR(X+Y) \leq VaR(X) + VaR(Y)$
  - VaR in Gaussian case has this peroperty, in other VaR, this may be violated
- Coherent Risk Measure:
  - Monotonicity: $Y \geq X \implies \rho(Y) \leq \rho(X)$, i.e. if Y has better value under almost all scenarios, then Y should have less risk
  - Subadditivity: $\rho(X+Y) \leq \rho(X) + \rho(Y)$
  - Positive-homogeneity: $\rho(\lambda X) = \lambda \rho(X)$, if $\lambda > 0$
  - Translation invariance: $\rho(X+c) = \rho(X) - c$
- ES and the worst-case analysis as in the SPAN system, are coherent risk measures
- VaR is coherent only under special assumptions about the distribution of returns (i.e. elliptical distribution)
- Proof of VaR is sub-additive under Gaussian case ...

7 Hot Spot Measuring Risk Contribution

- Marginal Risk:
  - MRisk is the change in the portfolio risk from taking an additional dollar of exposure to a given component:
  - $MRisk = \frac{\partial \rho_\alpha(w)}{\partial w}$
- Incremental Risk:
  - IRisk is the change in the portfolio risk owing to a new position from taking an additional dollar of exposure to a given component:
  - $IRisk = \rho_\alpha(w + \Delta w) - \rho_\alpha(w) = MRisk' \cdot \Delta w$
- Component Risk:
  - CRisk indicates how much the portfolio risk would change if the given component was deleted;
  - CRisk is the contribution of each component to the portfolio risk:
  - $CRisk = w. \times MRisk$
  - the sum of Component Risk returns the total portfolio risk (if the risk measure is a homogeneous risk function)
- In the Gaussian setting:

$$\begin{aligned} MVaR &= \frac{\partial VaR_\alpha(t, t+n\Delta)}{\partial w} = \mu_\Delta n - Z_{1-\alpha} \frac{\sum_\Delta \mathbf{w}}{\sqrt{\mathbf{w}' \sum_\Delta \mathbf{w}}} \sqrt{n} \\ CVaR &= \mathbf{w}. \times MVaR \\ VaR &= \mathbf{w}' MVaR \end{aligned} \tag{5}$$

- use EWMA to update VCV matrix

- Best hedge:
  - the VaR reduction given this hedge can be estimated as:
  - $\Delta VaR = MVaR_i \times \Delta w_i$
  - $\Delta w_i = -\frac{\sigma_{w,i}}{\sigma_i^2}$

8 Portfolio Modelling

- In the bottom up approach:
  - assgin the joint distribution of log-returns of different stocks
  - obtain the portfolio distribution
  - compute portfolio risk measures
- however, it is difficult to assign the joint distribution, particularly when the number of assets is large and the sample size is limited
- it is feasible to use multivariaet Gaussian or historical simulartion
- the exact portfolio log-return:

$$r_p(t, t+\Delta) = \log(\sum_{i=1}^{N} w_i e^{r_i(t,t+\Delta)}) = \log(\mathbf{w}' e^{\mathbf{r}(t,t+\Delta)}) \tag{6}$$

- but the distribution of r_p is not known in closed form, even if the log-returns are jointlt Gaussain
- we can use RM approximation:

$$r_p(t, t+\Delta) \simeq \sum_{i=1}^{N} w_i r_i(t) = \mathbf{w}'\mathbf{r}(t) \tag{7}$$

- use $e^x \simeq 1 + x$ and $\log(x) \simeq x$ to derive RM approximation
- RM approximation does not work over long horizon and negative weights
- The Parametric Gaussian Approach:
  - use RM approximation, we can have (under zero serial correlation):
    - $\mathbb{E}(r_p(t, t+\Delta)) = \mathbf{w}'\mu_\Delta$
    - $\mathbb{V}ar(r_p(t, t+\Delta)) = \mathbf{w}' \sum_\Delta \mathbf{w}$
-

$$\begin{aligned} VaR_\alpha^{P\&L}(t, t+n\Delta) &= P(t) \times (1 - e^{-VaR_\alpha^r(t,t+n\Delta)}) \\ ES_\alpha^{P\&L}(t, t+n\Delta) &\simeq P(t) \times ES_\alpha^r(t, t+n\Delta) \end{aligned} \tag{8}$$

- use MC simulation to simulate $r_p(t, t+\Delta)$, rather than using RM approximation:
  - assign the mean vector and the covariance matrix
  - simulate form a multivariate normal distribution (using the Cholesky decomposition) and a random vector $\mathbf{r}^{(i)}$
  - given the portfolio vectro $\mathbf{w}$, compute $r_p^{(i)} = \log(\mathbf{w}' e^{\mathbf{r}^{(i)}})$
  - repeat large number of times
  - compute risk measure
- Given the Cholesky matrix $\mathbf{A}$ and the VCV matrix $\mathbf{Z}$, the one-period simulated returns are: $\mathbf{r}(t) = \mathbf{AZ}(t)$, and simulated stock prices are: $\mathbf{P}(t+\Delta) = \mathbf{P}(t)e^{\mathbf{AZ}(t)}$

9 Value at Risk for Derivative Positions

- Exact formula:
  - possibel for a limited number of contracts for which there is a monotonic relation ship between risk factor and contract price
  - derivative price is an monotonic increasing function of the risk factor:
    * determine the worst risk factor scenario at the given confidence level: $P_{worst}(t+\Delta n) = P(t)e^{\mu\Delta n - Z_{1-\alpha}\sigma\Delta\sqrt{n}}$
    * revaluate the derivative position at the worst case scenario
  - derivative price is a monotonic decreasing function of the risk factor:
    * determine the best risk factor scenario at the given confidence level: $P_{best}(t+\Delta n) = P(t)e^{\mu\Delta n + Z_{1-\alpha}\sigma\Delta\sqrt{n}}$
    * revaluate the derivative position at the best case scenario
  - Limits:
    * the derivative price is easily computable
    * it is monotonic function of the underlying risk factor

- Full revaluation via MC simulation:
  - very general, and be abel to cope with accurate but very time consuming
  - Price the derivative position using the current value of the risk factor $P(t)$, i.e. compute $C(P(t), t)$
  - Simulate log-returns via our preferred model(either parametric Gaussian or historical simulation), so that we can obtain M simulated scenarios $r^i(t, t + n\Delta)$
  - Obtain the simulated risk-factor price at the VaR horizon t+nΔ, e.g. by $P^i(t, t+n\Delta) = P(t) \times e^{r^i(t,t+n\Delta)}$
  - Revaluate the derivative position at the time horizon under each simulated scenarios
  - Compute the M simulated $P\&L$ on the derivative position
  - Obtain VaR
  - Limits:
    * the revaluation step can be very costly (MC of MC)
    * use Taylor approximation to replace the repricing step

- Linear(Delta) approximation: fast but inaccurate

$$\Delta C(P(t), t) \simeq \frac{\partial C(P(t), t)}{\partial t} dt + \frac{\partial C(P(t), t)}{\partial P} P(t) \frac{dP(t)}{P(t)} \tag{9}$$

  - Delta-Normal VaR:
    * assume that percentage changes in the risk factor, i.e. $\frac{dP}{P}$, have a Gaussian distribution

$$VaR_{\alpha}^{delta-normal}(t, t+ndt) = -(\frac{\partial C(P(t), t)}{\partial t} ndt + \frac{\partial C(P(t), t)}{\partial P} P(t)\mu\Delta n + Z_{1-\alpha}|\frac{\partial C(P(t), t)}{\partial P})|P(t)\sigma\Delta\sqrt{n})$$

- Quadratic(Delta-Gamma) approximation: good tradeoff between accuracy and computational cost

$$\Delta C(P(t), t) \simeq \frac{\partial C(P(t), t)}{\partial t} dt + \frac{\partial C(P(t), t)}{\partial P} P(t) \frac{dP(t)}{P(t)} + \frac{1}{2} \frac{\partial^2 C(P(t), t)}{\partial P^2} P^2(t) (\frac{dP(t)}{P(t)})^2 \tag{10}$$

- finite difference method to approximate Greeks (but inaccurate)
- Delta-Gamma representation is valid only for small changes
- extend to multivariate case (Cholesky decomposition involved)

10 Parametric Value at Risk: The EWMA Approach

- estimate volatility using sample standard deviation:
  - this is an unconditional estimator, and in principle it cannot react to market shocks
  - cannot capture time-variation in the volatility
- The Risk-Metrics EWMA Model:

$$\sigma^2(t, t+\Delta) = \lambda\sigma^2(t-\Delta, t) + (1-\lambda)r^2(t-\Delta, t) \tag{11}$$

  - lower value of λ give more weight to the most recent observation
  - larger value of λ give more persistence to the variance series
  - use MLE to estimate λ:

$$\begin{aligned} &\log\mathcal{L}(r_0, r_\Delta, ..., r_{(T-1)\Delta}|\lambda, \sigma_0^2) \\ &= \sum_{j=0}^{T-1}(-\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\sigma_{j\Delta}^2) - \frac{1}{2}(\frac{r_{j\Delta}}{\sigma_{j\Delta}})^2) \end{aligned} \tag{12}$$

- Simulate returns in the EWMA model:
  - the simulated cumulative return is: $r(t, t+n\Delta) = \sum_{i=0}^{n-1} r_{\Delta,i}$
  - at each time step:
    * draw, independently from the previous extractions, a new random number ϵ_i according to a standard normal r.v.
    * use this random draw and the variance computed accodring to the EWMA scheme to simulate one-period log-return:

$$r_\Delta(t+i\Delta, t+(i+1)\Delta) = \sqrt{\hat{\sigma}^2(t+i\Delta, t+(i+1)\Delta)}\epsilon_i \tag{13}$$

    * the simulated cumulative retun is obtained by the sum of the simulated one-period log-returns
    * but n-periods log returns are not Gaussian
- Value at Risk with EWMA model
  - estimate the model
  - simulate M future scenarios to obtain M simulated values of $r(t, t+n\Delta)$
  - Compute the P&L distribution
  - Compute the empirical percentile of the M simulated P&L
- the multivariate extension:

$$\sum^{EWMA}(t, t+\Delta) = \lambda \sum^{EWMA}(t-\Delta, t) + (1-\lambda)\mathbf{r}(t-\Delta, t)\mathbf{r}'(t-\Delta, t) \tag{14}$$

- Attractions and limits of EWMA
  - Pros:
    * relatively little data needs to be stored (only current estimate of the variance and the most recent observation of return)
  - Cons:
    * does not allow for a leverage effect(i.e. negative dependency between return and volatility)
    * this problem can be solved by GARCH or EGARCH
- GARCH vs EWMA
  - EWMA has single parameter to control both market reaction and volatility persistence
  - GARCH(1,1) separates two effects:

$$\sigma^2(t) = w + \alpha r_{t-\Delta}^2(1 - \mathbb{1}\{r_{t-\Delta} > 0\}) + \gamma r_{t-\Delta}^2\mathbb{1}\{r_{t-\Delta} > 0\} + \beta\sigma_{t-\Delta}^2 \tag{15}$$

- the above equation is a GJR model that is derived from GARCH(1,1), which captures the leverage effect if $\alpha \neq \gamma$

- Stochastic Volatility Models:
  - in SV models either the conditional and unconditional variance are assumed to be stochastic
  - widely used in option pricing(can derive closed form solutions via Fourier transform)
  - not used in risk-management due to difficulties in constructing the likelihood function
  - Log-price dynamics: $d\text{logP(t)} = \mu dt + \sqrt{v(t)}dW_1(t)$
    Instantaneous variance dynamics: $dv(t) = \alpha(v(t),t)dt + \beta(v(t),t)dW_2(t)$ (Hull& White model)
  - Log-price dynamics: $d\text{logP(t)} = \mu dt + \sqrt{v(t)}dW_1(t)$
    Instantaneous variance dynamics: $dv(t) = k(\theta - v(t))dt + \epsilon\sqrt{v(t)}dW_2(t)$ (Heston model)
    * mean-reverting
    * positive volatility

11 Multifactor Models, Dimension Reduction and Principal Component Analysis

- PCA gives the posibility to identify volattility factors from a time series of historical term structure
- $\sum$ is the covariance matrix
- eigenvalues are solutions of the linear system: $\det(\sum - \lambda \mathbf{I_2}) = 0$
- the sum of eigenvalues is equal to trace($\sum$)
- normalised eigenvectors at a given eigenvalue are solutions of the linear system:

$$(\sum - \lambda_i \mathbf{I_2})\mathbf{x}_{\lambda_i} = \mathbf{0}_2 \tag{16}$$

- the i-th factor accounts for $\frac{\lambda_i}{\text{trace}(\sum)}$ of the total variance