L3 Portfolio Optimization

Premier on Financial Data

log-prices $y_t \triangleq \log P_t$. follows random walk $y_t = \mu + y_{t-1} + \varepsilon_t$

simple return $R_t \triangleq \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{P_t}{P_{t-1}} - 1$

log-return $r_t \triangleq y_t - y_{t-1} = \log\left(\frac{P_t}{P_{t-1}}\right) = \log(1 + R_t)$ 捕捉price差分

log-returns: diff(log(prices))[-1]

linear (simple) return: (prices/lag(prices)-1)[-1]

why log-return?

① prices are assumed to be distributed log normally $\Rightarrow r_t$ normally distributed

② $r_t \approx R_t$ when returns are small. good approximation

③ Compound return: $\prod_i (1+R_i) = (1+R_1)(1+R_2)\cdots(1+R_i)$

$\log \prod_i (1+R_i) = r_1 + r_2 + \cdots + r_i \rightarrow$ time-additivity. normally distributed

i.i.d model

$$\boxed{r_t = \mu + w_t}$$

parameter estimation: $\mu \triangleq E[r_t \mid F_{t-1}] = \hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} r_t$ (future; sample mean)

$$\Sigma \triangleq \mathrm{Cov}[r_t \mid F_{t-1}] = E[(r_t - \mu_t)(r_t - \mu_t)^T \mid F_{t-1}]$$

$$= \hat{\Sigma} = \frac{1}{T-1}\sum_{t=1}^{T}(r_t - \hat{\mu})(r_t - \hat{\mu})^T$$

sample covariance matrix

∵ contains too much estimation error, so a portfolio design based on these estimates (Markowitz's portfolio) can be severely affected.

Portfolio Basics

portfolio w: $1^T w = 1$

portfolio return: $R_t^P = \sum_{i=1}^{N} w_i(R_{it}+1) - 1 = \sum_{i=1}^{N} w_i R_{it} \approx \sum_{i=1}^{N} w_i r_{it} = w^T r_t$

L3 cont'd

Performance Measures

Expected return : $w^T\mu$; variance : $w^T\Sigma w$

Volatility : $\sqrt{w^T\Sigma w}$

Sharpe Ratio (SR) : $\dfrac{w^T\mu - r_f}{\sqrt{w^T\Sigma w}}$ ← risk-free rate

Information Ratio (IR) : SR with respect to a benchmark $= \dfrac{E[w^T r_t - r_{b,t}]}{\sqrt{Var[w^T r_t - r_{b,t}]}}$

Drawdown : decline from a historical peak of the cumulative profit $X(t)$

$$D(T) = \max_{t\in[0,T]} X(t) - X(T)$$

VaR (Value at Risk) : quantile of the loss

ES (Expected Shortfall) or CVaR (Conditional Value at Risk) :
expected value of the loss above some quantile

Practical Constraints

Capital budget : $1^T w = 1$

Long-only : $w \geq 0$

Holding : $L \leq w \leq U$ (diversify)

Leverage : $\|w\|_1 \leq L$

Cardinality : $\|w\|_0 \leq K$

Turnover : $\|w - w_0\|_1 \leq u$. w_0 is the current portfolio

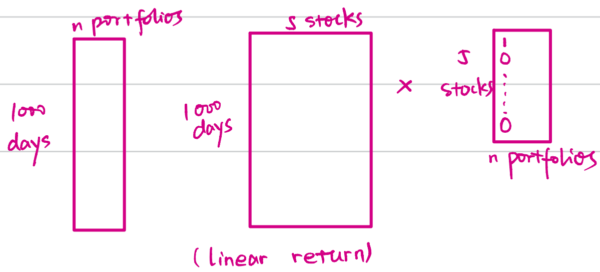
Market-neutral : $\beta^T w = 0$

Heuristic Portfolios

Buy & Hold (B&H)

$w = e_i$

return = X_lin %*% w



1 + cumsum(return) simple return

cumprod(1+ return) compound (geometric) return

(exp. L3.html page 37)

L3 cont'd

Equally weighted portfolio (EWP) or 1/N portfolio $w = \frac{1}{N}\mathbf{1}$

Quantile Portfolio : rank by μ ; $\frac{\mu}{\text{diag}(\Sigma)}$; $\frac{\mu}{\sqrt{\text{diag}(\Sigma)}}$

Global maximum return portfolio (GMRP)
$$\begin{bmatrix} \underset{w}{\text{maximize}} & w^T\mu \\ \text{s.t.} & 1^T w = 1,\ w \geq 0 \end{bmatrix}$$

all in the "best" stock

"Not" Heuristic Portfolios

Markowitz's mean-variance portfolio

$$\begin{bmatrix} \underset{w}{\text{maximize}} & w^T\mu - \lambda w^T\Sigma w \\ \text{s.t.} & 1^T w = 1 \\ & \vdots \ \text{constraints} \end{bmatrix}$$

Global minimum variance portfolio (GMVP)

$$\begin{bmatrix} \underset{w}{\text{minimize}} & w^T\Sigma w \\ \text{s.t.} & \mathbf{1}^T w = \mathbf{1} \end{bmatrix}$$

Maximum Sharpe ratio portfolio (MSRP)

$$\begin{bmatrix} \underset{w}{\text{maximize}} & \dfrac{w^T\mu - r_f}{\sqrt{w^T\Sigma w}} \\ \text{s.t.} & \mathbf{1}^T w = \mathbf{1} \end{bmatrix}$$ nonconvex

Risk-based portfolios

Inverse volatility portfolio (IVP)

$$w = \frac{\sigma^{-1}}{1^T\sigma^{-1}} \quad \text{where } \sigma^2 = \text{Diag}(\Sigma)$$

L3 cont'd

(MSRP) $\mu \to \sigma$

Most diversified portfolio (MDP)

$$\left[\begin{array}{ll} \underset{w}{\text{maximize}} & \underbrace{\dfrac{w^T\sigma}{\sqrt{w^T\Sigma w}}}_{\text{diversification ratio}} \\ \text{s.t.} & 1^T w = 1 \end{array}\right]$$

diversification ratio

Maximum decorrelation portfolio (MDCP)

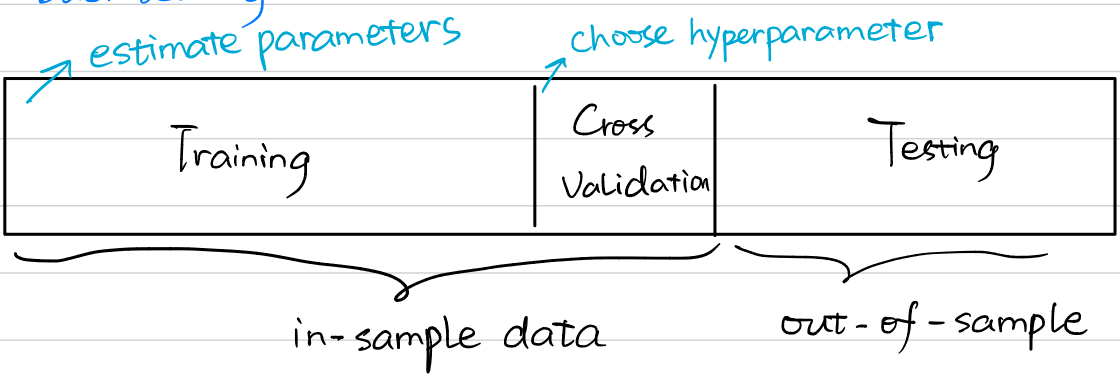
(GMVP) $\Sigma \to C$

$$\left[\begin{array}{ll} \underset{w}{\text{maximize}} & w^T C w \\ \text{s.t.} & 1^T w = 1 \end{array}\right] \qquad C = \mathrm{Diag}(\Sigma)^{-1/2}\,\Sigma\,\mathrm{Diag}(\Sigma)^{-1/2}$$

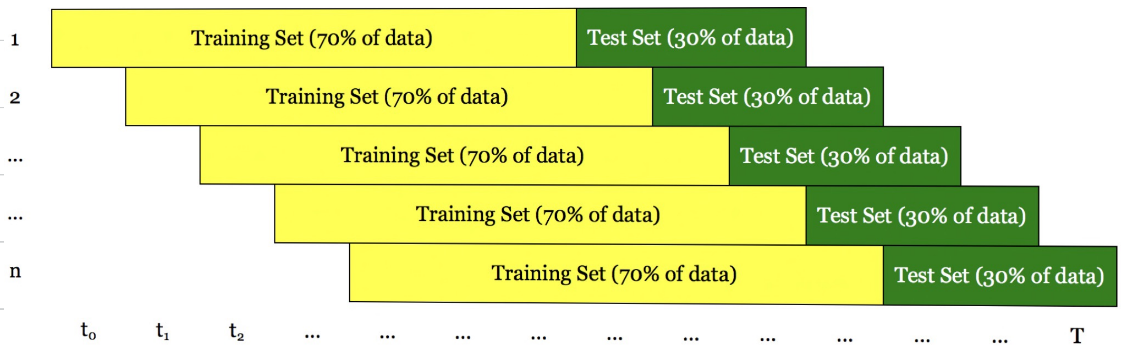
L4 Backtesting

Backtesting with Historical Market Data

Vanilla backtesting



Walk-forward backtesting



Multiple Randomized Backtesting

perform multiple backtests of portfolios in an automated way on a rolling-window basis by taking data randomly from different markets, different time periods, and different stock universes.

Backtesting with synthetic data

Monte Carlo simulations:

① resampling the existing history: e.g. sample the realized sequence and even reorder; although it isn't iid, there are solutions

② creating a synthetic dataset: data → model → new data

L5 : Shrinkage and Black-Litterman

Shrinkage estimator

Intuition : Lower estimation errors can be achieved by allowing

small sample   some bias in exchange of a smaller variance

$$\hat{\theta}^{sh} = (1-\underbrace{\rho}_{\text{shrinkage trade-off par}})\hat{\theta} + \rho \underbrace{\theta^{target}}_{\text{prior information}}$$

△ choice of ρ is essential $\begin{cases} \text{cross-validation} \\ \text{Random Matrix Theory (RMT)} \end{cases}$

Intepretation: Linear shrinkage of the eigenvalues (while keeping the same eigenvectors) towards one :

$$\lambda_i(\hat{\Sigma}^{sh}) = (1-\rho)\lambda_i(\hat{\Sigma}) + \rho 1$$

Shrinkage for the Mean

sample mean estimator : $\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} r_t$, $\hat{\mu} \sim N(\mu, \frac{1}{T}\Sigma)$

MSE : $E[\|\hat{\mu}-\mu\|^2] = \frac{1}{T}Tr(\Sigma)$

shrinkage

James-Stein Estimator : $\left(\hat{\mu}^{JS} = (1-\rho)\hat{\mu} + \rho t\right) \rightarrow E[\|\hat{\mu}^{JS}-\mu\|^2] \leq E[\|\hat{\mu}-\mu\|^2]$

$$\rho = \frac{1}{T}\frac{N\bar{\lambda} - 2\lambda_{max}}{\|\hat{\mu}-t\|^2} ,$$ where $\bar{\lambda}$ and λ_{max} (eigenvalues of Σ)

L5: cont'd

Shrinkage for the Covariance Matrix

$$\left(\hat{\Sigma}^{sh} = \underline{(1-\rho)} \hat{\Sigma} + \underline{\rho} T \right)$$

实际不一定需要和为1

to find ρ

Method 1:

weird: we don't know true Σ (RMT comes)

$$\left[\begin{array}{ll} \underset{\rho_1, \rho_2}{\text{minimize}} & E[\| \hat{\Sigma}^{sh} - \Sigma \|_F^2] \\ \text{s.t.} & \hat{\Sigma}^{sh} = \rho_1 I + \rho_2 \hat{\Sigma} \end{array} \right]$$

optimization problem

RMT: for large T and N, can derive an estimator without knowledge of Σ

Ledoit-Wolf:

$$\rho = \min\left(1, \frac{\frac{1}{T^2} \sum_{t=1}^{T} \| \hat{\Sigma} - r_t r_t^T \|_F^2}{\| \hat{\Sigma} - T \|_F^2}\right)$$

Method 2: Maximizing the Sharpe Ratio

let $\hat{\Sigma}^{sh} = \rho_1 I + \hat{\Sigma}$

Optimal portfolio $w^{SR} \propto (\hat{\Sigma}^{sh})^{-1} \hat{\mu}$

$$\left[\begin{array}{ll} \underset{\rho_1 \geq 0}{\text{maximize}} & \dfrac{\hat{\mu}^T (\hat{\Sigma}^{sh})^{-1} \mu}{\sqrt{\hat{\mu}^T (\hat{\Sigma}^{sh})^{-1} \Sigma (\hat{\Sigma}^{sh})^{-1} \hat{\mu}}} \\ \text{s.t.} & \hat{\Sigma}^{sh} = \rho_1 I + \hat{\Sigma} \end{array} \right] \rightarrow \text{RMT}$$

L5: cont'd

Black - Litterman Model

Intuition: Put together 2 sources of information:
market equilibrium + investor's view

Market Equilibrium:

Choice 1: sample mean $\hat{u} = \frac{1}{T}\sum_{t=1}^{T} r_t$, the estimate $\pi = \hat{\mu}$

can writes in terms of the actual u and estimation err.

$$\left(\pi = \mu + w \ , \quad w \sim \mathcal{N}(0, \tau\Sigma) \right) \quad ①$$

unknown

Choice 2: CAPM

Investor's View:

suppose we have k views in the form of:

$$\left(v = P\mu + e \ , \quad e \sim \mathcal{N}(0, \Omega) \right) \quad ②$$

$$\begin{cases} P \in R^{k\times N} \text{ 加上 } v \in R^{k} : \text{刻画 absolute or relative } k \text{ views} \\ \Omega \in R^{k\times k} : \text{ measure the uncertainty in the views} \end{cases}$$

Black - Litterman Model - Weighted LS Approach

combine ① and ②:

$$\left(y = X\mu + \epsilon \ , \quad \epsilon \sim \mathcal{N}(0, V) \right)$$

$$y = \begin{bmatrix} \pi \\ v \end{bmatrix}, \quad X = \begin{bmatrix} I \\ P \end{bmatrix}, \quad V = \begin{bmatrix} \tau\Sigma & 0 \\ 0 & \Omega \end{bmatrix}$$

$\downarrow$

weighted LS: $\left[\underset{\mu}{\text{minimize}} \ (y - X\mu)^T V^{-1} (y - X\mu) \right]$

solution: $\hat{\mu}_{BL} = W_{mkt}\hat{\mu}_{mkt} + W_{views}\hat{\mu}_{views}$

note that: $W_{mkt} + W_{views} = I$

L6 - Estimators under heavy tails and outliers

Motivation

① Data has heavy tails, but we assume gaussian $N(\hat{u}, \hat{\Sigma})$ ✓ heavy-tail distribution

② Data is gaussian, but has outliers

example: Student's t-distribution $v \uparrow$, heavy tails $\downarrow$

advantage of robust estimators: assume heavy tails

if heavy tails are true (i.e. v is small in real data), then improve a lot

otherwises if doesn't have heavy tails, it won't pay much. (although sample mean & sample covariance are optimal

Robust Covariance Matrix Estimators (假设已知μ)

Robust M-estimator (General)

Intuition: different weights for outliers

recall MLE for gaussian

$$f(x) = C \det(\Sigma)^{-\frac{1}{2}} \exp(-\tfrac{1}{2} x^T \Sigma^{-1} x)$$

neg-log: $L(\Sigma) = \frac{N}{2} \log\det(\Sigma) + \sum_{i=1}^{N} x^T \Sigma^{-1} x$

$\rightarrow$ $\hat{\Sigma} = \frac{1}{N} \sum_{i=1}^{N} x_i x_i^T$ (sample covariance)

M-estimator

$$L(\Sigma) = \frac{N}{2} \log\det(\Sigma) + \sum_{i=1}^{N} \rho(x_i^T \Sigma^{-1} x_i)$$ add ρ function

$\rightarrow$ $\Sigma = \frac{1}{N} \sum_{i=1}^{N} \underbrace{w(x_i^T \Sigma^{-1} x_i)}_{\text{weight}} x_i x_i^T$ $\quad w = \frac{\rho'}{2}$ if differentiable

problem: Σ appears in both sides (called fixed-point equation)

L6 - cont'd

Tyler's M-estimator for Elliptical Distributions (special case)

2 key components of the distribution: shape + tail

Intuition: by normalization, it doesn't matter how heavy the tail is

By normalization: $s_i \triangleq \frac{x_i}{\|x_i\|_2}$ no tail

pdf: $f(s) = C \det(R)^{-\frac{1}{2}} (s^T R^{-1} s)^{-\frac{k}{2}}$ $R \sim \Sigma$ is covariance matrix. shape preserves

MLE on $f(s)$ $\to$ $\mathcal{L}(\Sigma) = \frac{N}{2} \log\det(\Sigma) + \frac{k}{2} \sum_{i=1}^{N} \underbrace{\log(s_i^T \Sigma^{-1} s_i)}_{x_i^T \Sigma^{-1} x_i}$ log is a $\rho(\cdot)$

$$\to \quad \Sigma = \sum_{i=1}^{N} w_i x_i x_i^T, \quad w_i = \frac{k}{N(x_i^T \Sigma^{-1} x_i)}$$

weights $\propto 1/d_i^2$, outliers are down-weighted

interative algorithm to solve fixed-point equation

$$\left(\begin{aligned} \tilde{\Sigma}_{t+1} &= \frac{k}{N} \sum_{i=1}^{N} \frac{x_i x_i^T}{x_i^T \Sigma^{-1} x_i} \\ \Sigma_{t+1} &= \tilde{\Sigma}_{t+1} / \mathrm{Tr}(\tilde{\Sigma}_{t+1}) \end{aligned} \right)$$

Robust Mean-Covariance Estimators (假设未知 μ)

Robust M-estimators for location and scatter (general)

- for μ: $\frac{1}{N} \sum_{i=1}^{N} \underbrace{u_1((x_i-\mu)^T R^{-1} (x_i-\mu))}_{\text{image} = 1,\ \mu \text{ becomes sample mean}} (x_i - \mu) = 0$

weights

- for Σ: $\frac{1}{N} \sum_{i=1}^{N} u_2((x_i-\mu)^T R^{-1} (x_i-\mu)) (x_i - \mu) = R$

MLE for Student's t-distribution:

$$f(x) = C \det(R)^{-\frac{1}{2}} \left(1 + \frac{1}{v} (x-\mu)^T R^{-1} (x-\mu)\right)^{-\frac{k+v}{2}}$$

$$\mathcal{L}(\mu, R) = \frac{N}{2} \log\det(R) + \frac{k+v}{2} \sum_{i=1}^{N} \log(v + (x_i-\mu)^T R^{-1} (x_i-\mu))$$

$$\Rightarrow \frac{k+v}{N} \sum_{i=1}^{N} \frac{x_i - \mu}{v + (x_i-\mu)^T R^{-1} (x_i-\mu)} = 0, \quad \frac{k+v}{N} \sum_{i=1}^{N} \frac{(x_i-\mu)(x_i-\mu)^T}{v + (x_i-\mu)^T R^{-1} (x_i-\mu)} = R$$

L6 - cont'd

Small Sample Regime

Shrinkage Robust Estimator with Known Mean

Wiesel's penalty: $h(\Sigma) = \log\det(\Sigma) + K \log \mathrm{Tr}(\Sigma^{-1} \underset{\text{targeted}}{\underbrace{T}})$

Penalized loss function:

$$\mathcal{L}(\Sigma) = \frac{N}{2} \log\det(\Sigma) + \frac{k}{2} \sum_{i=1}^{N} \log(x_i^T \Sigma^{-1} x_i) + \alpha\left(\log\det(\Sigma) + k \log \mathrm{Tr}(\Sigma^{-1} T)\right)$$

Algorithm 1:

$$\Sigma_{t+1} = \frac{N}{N+2\alpha} \underbrace{\frac{k}{N} \sum_{i=1}^{N} \frac{x_i x_i^T}{x_i^T \Sigma_t^{-1} x_i}}_{\approx \text{ Tyler}} + \frac{2\alpha}{N+2\alpha} \frac{KT}{\mathrm{Tr}(\Sigma_t^{-1} T)}$$

Shrinkage closer to target

Algorithm 2: MM of $\mathcal{L}(\Sigma)$

no $p(\cdot)$, more like gaussian

$$g(\Sigma \mid \Sigma_t) = \frac{N}{2} \log\det(\Sigma) + \frac{k}{2} \sum_{i=1}^{k} \frac{x_i^T \Sigma^{-1} x_i}{x_i^T \Sigma_t^{-1} x_i} + \alpha\left(\log\det(\Sigma) + k \frac{\mathrm{Tr}(\Sigma^{-1} T)}{\mathrm{Tr}(\Sigma_t^{-1} T)}\right)$$

update:

$$\tilde{\Sigma}_{t+1} = \frac{N}{N+2\alpha} \frac{k}{N} \sum_{i=1}^{N} \frac{x_i x_i^T}{x_i^T \Sigma_t^{-1} x_i} + \frac{2\alpha}{N+2\alpha} \frac{KT}{\mathrm{Tr}(\Sigma_t^{-1} T)}$$

normalization:

$$\tilde{\Sigma}_{t+1} = \tilde{\Sigma}_{t+1} / \mathrm{Tr}(\tilde{\Sigma}_{t+1})$$

coverages to unique solution

Shrinkage Robust Estimator with Unknown Mean

brings together robust estimators of mean & covariance matrix and shrinkages of mean & covariance matrix

4] Robust Portfolio (copy from convex optimization)

Robust Optimization

Problem formulation: some problem contains parameters θ that are typically estimated in practice $(\hat{\theta}) \Rightarrow x^*(\hat{\theta}) \neq x^*(\theta)$

Several ways to make the problem robust to parameters errors:

manageable
① stochastic robust optimization (involving expectations) 假设有概率分布(普遍情况)
② worst-case robust optimization 太悲观
③ chance programming or chance robust optimization best but difficult (nonconvex)

Stochastic robust optimization: Expectations

Instead of using approximated function $f(x;\hat{\theta})$, use $E_\theta[f(x;\theta)]$.

e.g. ① model the estimated value as $\theta = \hat{\theta} + \delta$, $\delta \sim N(0, \sigma)$

② consider $f(x;\hat{\theta}) = (\hat{c}^T x)^2$

$$\rightarrow E_\theta[f(x;\theta)] = E_\delta[((\hat{c}+\delta)^T x)^2]$$
$$= E_\delta[x^T \hat{c}\hat{c}^T x + x^T \delta\delta^T x]$$
$$= (\hat{c}^T x)^2 + \underbrace{x^T Q x}_{\text{regularizer}}$$

Convexity: if $f(x;\theta)$ convex, then expectation (nonnegative sum) is convex

Worst-case robust optimization

Assume that true parameter lies in an uncertainty region centered around the estimated value: $\theta \in \mathcal{U}$

关键: ① $\mathcal{U}$ 好解 ② $\mathcal{U}$ 不大

sephere region: $\mathcal{U} = \{\theta \mid \|\theta - \hat{\theta}\|_2 \leq \delta\}$

box region: $\mathcal{U} = \{\theta \mid \|\theta - \hat{\theta}\|_\infty \leq \delta\}$

elliptical region: $\mathcal{U} = \{\theta \mid (\theta - \hat{\theta})^T S^{-1} (\theta - \hat{\theta}) \leq \delta^2\}$

L7 cont'd (copy from convex optimization)

e.g. consider a sphere uncertainty region $\mathcal{U} = \{c \mid \|c-\hat{c}\|_2 \le \delta\}$

consider $f(x;\hat{\theta}) = (\hat{c}^T x)^2$ if the function is the objective to be minimized (maximized) or it is a constraint of the form $f(x;\hat{\theta}) \le 0$ ($\ge$)

$\rightarrow$ worst case: $\max_{c\in\mathcal{U}} |c^T x| = \max_{\|e\|\le\delta} |(\hat{c}+e)^T x|$ sephere

Cauchy-Schwarz inequality $|u^T v| \le \|u\|\|v\|$ equal iff $u = \lambda v$

$\le \max_{\|e\|\le\delta} (|\hat{c}^T x| + |e^T x|)$

$\le |\hat{c}^T x| + \delta \|x\|$ ← the upper bound can achieve (maximum)

要minimize的objective，考虑它worst case则是 min max

要是 constraint ≤ 0，考虑它的worst case是 max ≤ 0

一般要求max有闭式解，否则问题会复杂

convexity: if C is convex set, f is convex, then pointwise maximum is convex ← 系列不同C的function

加上robustness，形式上会变复杂（例如 LP/QP → SOCP → SDP）

① Robust Global Maximum Return Portfolio Optimization

$$\max_w \; w^T\mu \quad \text{s.t.} \; 1^T w = 1 \quad \Rightarrow \quad \max_w \min_{\mu\in\mathcal{U}_\mu} w^T\mu \quad \text{s.t.} \; 1^T w = 1 \quad \Rightarrow \quad \max_w \; w^T\hat{\mu} - \kappa\|S^{1/2}w\|_2 \quad \text{s.t.} \; 1^T w = 1$$

LP

add regularizer

SOCP (epigraph form)

(assume $\mathcal{U}_\mu = \{\mu = \hat{\mu} + \kappa S^{1/2} u \mid \|u\|_2 \le 1\}$)

ellipsoid. $S^{1/2}$ is a square root of S ($\because (S^{1/2})^T S^{1/2} = S$). note: not unique considering $Q\cdot S^{1/2}$

"Cholesky decomposition"

② Robust Global Minimum Variance Portfolio Optimization

$$\min_w \; w^T\Sigma w \quad \text{s.t.} \; 1^T w = 1 \quad \Rightarrow \quad \min_w \max_{\Sigma\in\mathcal{U}_\Sigma} w^T\Sigma w \quad \text{s.t.} \; 1^T w = 1 \quad \Rightarrow \cdots \Rightarrow \quad \min_w \; \underbrace{\|\hat{X}w\|_2}_{\le t_1} + \underbrace{\delta_x\|w\|_2}_{\le t_2} \quad \text{s.t.} \; 1^T w = 1$$

QP

SOCP (epigraph form)

↓

Matrix version robustness

$\hat{\Sigma} = \frac{1}{T} X^T X$, $X = \hat{X} + \Delta$, $\mathcal{U}_X = \{X \mid \|X - \hat{X}\|_F \le \delta_X\}$

"Tikhonov regularization"

QP 变形，都加上平方（效果要试了才知道）

$$\min_w \; \|\hat{X}w\|_2^2 + \delta_x\|w\|_2^2 = w^T(\hat{X}^T\hat{X} + \delta_x I)w \quad \text{s.t.} \; 1^T w = 1$$

looks like sample covariance matrix + shrinkage to I

L8 Portfolio Optimization with Alternative Risk Measures

Markowitz portfolio (recall)

$$\begin{bmatrix} \underset{w}{\text{maximize}} & w^T\mu - w^T\Sigma w \\ \text{s.t.} & 1^T w = 1 \end{bmatrix}$$

Intuition 1: use different risk measures

Intuition 2: we don't use μ and Σ in this chapter

Alternative Measures of Risk

Downside risk: only penalize downside variance

- semi-variance: $SV = E[((E[R] - R)^+)^2]$
- lower partial moments (LPM): $LPM = E[((\tau - R)^+)^\alpha]$

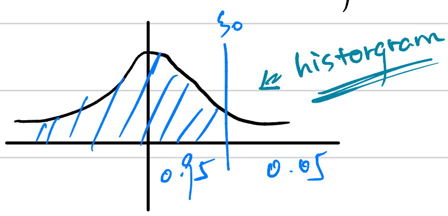
general

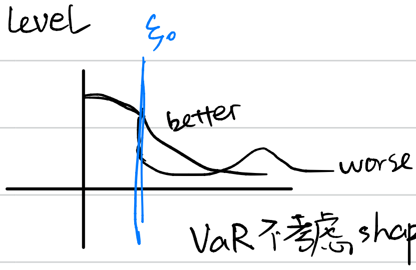
Value-at-Risk (VaR): maximum loss with a specified confidence level

$$VaR_\alpha = \inf\{\xi_0 : \Pr(\xi \le \xi_0) \ge \alpha\} \qquad \text{loss } \xi = -w^T r$$

confidence level

e.g. $\alpha = 0.95$





VaR不考虑 shape of tail

improvement

Conditional Value-at-Risk (CVaR):

account shape of the losses exceeding the VaR

$$CVaR_\alpha = E[\xi \mid \xi \ge VaR_\alpha]$$

often requires large # sample because the ratio of tail is small

Drawdown (DD)

unnormalized version: $D(t) = \max_{1 \le \tau \le t} X(\tau) - X(t)$

normalized version: $D(t) = \dfrac{\max_{1 \le \tau \le t} X(\tau) - X(t)}{\max_{1 \le \tau \le t} X(\tau)}$

time series

↓

max / ave / VaR on DD one number

上述计算容易，设计 portfolio 去 minimize 则复杂

L8 cont'd

Mean-semi-variance portfolio

$$\left[\begin{array}{ll} \underset{w}{\text{maximize}} & w^T\mu - \lambda \frac{1}{T}\sum_{t=1}^{T} \underbrace{\left((w^T\mu - w^T r_t)^+\right)^2}_{=t} \\ \text{s.t.} & 1^T w = 1 \ , \ w \geq 0 \end{array}\right]$$

convex $\downarrow$ QP

△ trick to remove $(\cdot)^+$:

constraints: $t = \max(0, \cdot\cdot) \Rightarrow t \geq \max(0, \cdot\cdot) \Rightarrow \begin{cases} t \geq 0 \\ t \geq \cdots \end{cases}$

Markowitz portfolio has a similar form:

$$\left[\begin{array}{ll} \underset{w}{\text{maximize}} & w^T\mu - \lambda \frac{1}{T}\sum_{t=1}^{T} (w^T\mu - w^T r_t)^2 \\ \text{s.t.} & 1^T w = 1 \ , \ w \geq 0 \end{array}\right]$$

Mean-CVaR portfolio

$$\left[\begin{array}{ll} \underset{w}{\text{maximize}} & w^T\mu \\ \text{s.t.} & \text{CVaR}_\alpha(-w^T r) \leq c \\ & 1^T w = 1 \ , \ w \geq 0 \end{array}\right]$$

$\leftarrow$ put in constraint / objective

CVaR portfolio

$$\left[\begin{array}{ll} \underset{w}{\text{minimize}} & \text{CVaR}_\alpha(-w^T r) \\ \text{s.t.} & 1^T w = 1 \ . \ w \geq 0 \end{array}\right]$$

Difficulty: given w. we can easily gain VaR then CVaR; but we have to switch w

→ Solution:

nonneg. sum + convex → convex

auxiliary convex function: $F_\alpha(w, \zeta) = \zeta + \frac{1}{1-\alpha} E[\underbrace{-w^T r - \zeta}_{\text{linear}}]^+_{\text{max}}$

$$\begin{cases} \text{VaR}_\alpha(-w^T r) \in \text{argmin}_\zeta F_\alpha(w, \zeta) \\ \text{CVaR}_\alpha(-w^T r) \in \min_\zeta F_\alpha(w, \zeta) \end{cases}$$

△ using same trick to deal with $(\cdot)^+$

$$\left[\begin{array}{ll} \underset{w, \{z_t\}, \zeta}{\text{minimize}} & \zeta + \frac{1}{1-\alpha} \frac{1}{T}\sum_{t=1}^{T} z_t \\ \text{s.t.} & 0 \leq z_t \geq -w^T r_t - \zeta \ , \ t = 1, \cdots, T \\ & 1^T w = 1 \ , \ w \geq 0 \end{array}\right]$$

LP (but many variables)

L8 cont'd

Mean - DD portfolio

cumulative (uncompounded) return : $r^{cum}(t) = \sum_{\tau=1}^{t} r(\tau)$

portfolio cumulative return : $r_p^{cum}(\tau) = w^T r^{cum}(\tau)$

drawdown (DD) : $D(t) = \max_{1\le\tau\le t} r_p^{cum}(\tau) - r_p^{cum}(t)$ convex

Mean - Max - DD portfolio

$$\begin{bmatrix} \underset{w}{\text{maximize}} & w^T\mu \\ \text{s.t.} & \underbrace{\max_{1\le t\le T}}\left\{\underbrace{\max_{1\le\tau\le t}} w^T r_\tau^{cum} - w^T r_t^{cum}\right\} \le c \\ & 1^T w = 1, \quad w \ge 0 \end{bmatrix}$$

$$\Downarrow$$

$$\begin{bmatrix} \underset{w,\{u_t\}}{\text{maximize}} & w^T\mu \\ \text{s.t.} & u_t - w^T r_t^{cum} \le c, \quad \forall 1 \le t \le T \\ & u_t \ge w^T r_\tau^{cum}, \quad \forall 1\le t\le T, \ 1\le\tau\le t \\ & 1^T w = 1, \quad w \ge 0 \end{bmatrix}$$

$$\Downarrow$$

$$\begin{bmatrix} \underset{w,\{u_t\}}{\text{maximize}} & w^T\mu \\ \text{s.t.} & u_t - w^T r_t^{cum} \le c, \quad \forall 1 \le t \le T \\ & \begin{cases} u_t \ge w^T r_t^{cum} \\ u_t \ge u_{t-1} \end{cases} \\ & 1^T w = 1, \quad w \ge 0 \end{bmatrix}$$

LP

L11 time series modeling

Variance / Covariance Models – Volatility Clustering

Moving average (MA)

$$\sigma_t^2 = \frac{1}{m}\sum_{i=1}^{m} w_{t-i}^2 \quad \text{(square of return)}$$

实际上 return $r_t = \mu_t + \underbrace{w_t}_{\text{(residual)}}$. 但 μ_t 很小可忽略 .

Exponentially Weighted MA (EMMA)

$$\sigma_t^2 = \alpha w_{t-1}^2 + (1-\alpha)\sigma_{t-1}^2$$

ARCH model

Latent variable

$$w_t = \sigma_t Z_t$$

$$\sigma_t^2 = w + \sum_{i=1}^{m} a_i w_{t-i}^2$$

σ_t : envelope , changes slowly

Z_t : white noise , zero mean & constant unit variance

作用是 model the probability → param. estimation

similar to MA: increase m a lot for smoother output

GARCH model (Generalized ARCH)

$$w_t = \sigma_t Z_t$$

$$\sigma_t^2 = w + \sum_{i=1}^{m} a_i w_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2$$

ARMA : even (1, 1) can produce smooth output in e.g.

critisim : GARCH is a glorified exponential smoothing
(以指数衰减 刻画 envelope)

Stochastic Volatility Model

$$w_t = \sigma_t Z_t$$

$$\log(\sigma_t^2) = \bar{h} + \phi(\log(\sigma_{t-1}^2) - \bar{h}) + u_t$$