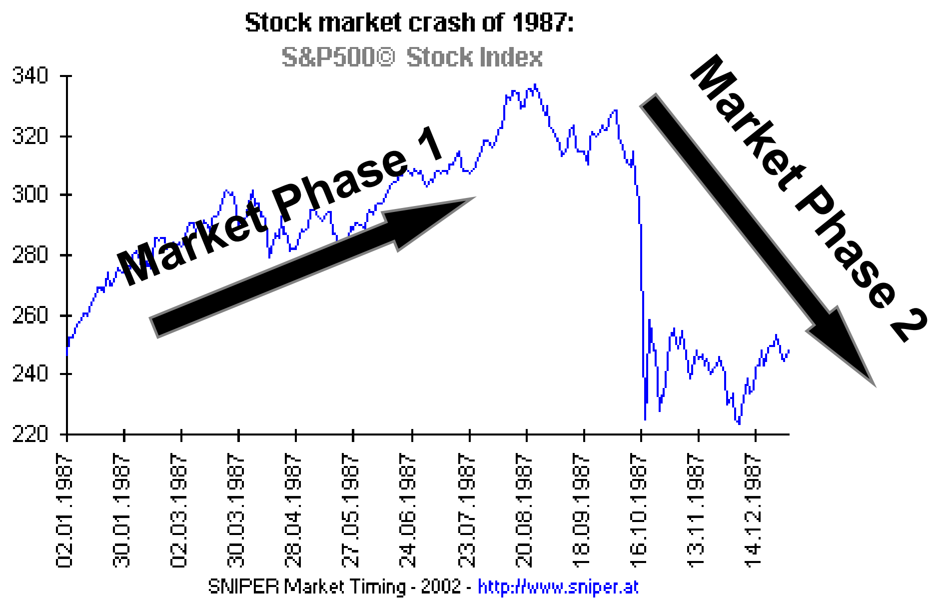
Estimation of multidimensional SDEs and Change-Point Detection

I. Horenko

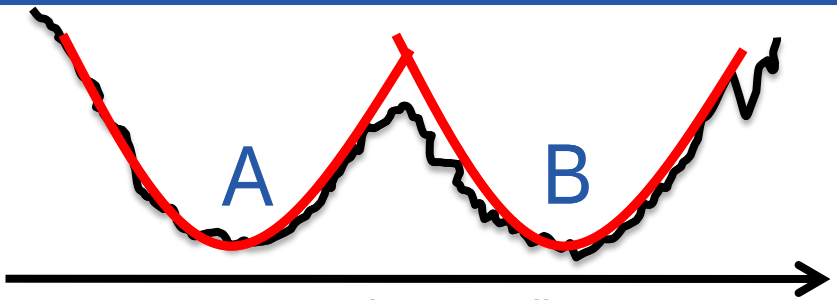
DFG Research Center MATHEON
Mathematics for key technologies

06.02.2009

Computational Finance



Idea: identify the change in the market by the change of the Black-Scholes model parameters



$$\dot{\boldsymbol{z}}(t) = F_{h(t)}\left(\boldsymbol{z} - \boldsymbol{\mu}_{h(t)}\right) + \Sigma_{h(t)}\dot{\mathbf{W}}(t) \qquad F_{h(t)}, \Sigma_{h(t)} \in \mathbb{R}^{n \times n}$$

$$h(t) \in \{1, \ldots, s\}, \qquad \boldsymbol{\mu}_{h(t)} \in \mathbb{R}^{n}$$

- An intuitive model for fluctuation around a stable conformation.
- Allows formulation in context of FEM-clustering (see previous lectures).

This leads to the following mathematical question:

Market Data → $Z = \{\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots \boldsymbol{z}_T\}$

Is there a $t \in \{t_1, \ldots, t_2\}$ such that

$Z_1 = \{\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots \boldsymbol{z}_t\}$ and $Z_2 = \{\boldsymbol{z}_{t+1}, \boldsymbol{z}_{t+2} \ldots \boldsymbol{z}_T\}$

are generated by linear SDE's with different parameters?

In fact, we arrived at a model selection problem:

Is a model with parameter space $\theta_1 = \{F_1, \Sigma_1, \boldsymbol{\mu}_1\}$ or with $\theta_2 = \{F_1, \Sigma_1, \boldsymbol{\mu}_1, F_2, \Sigma_2, \boldsymbol{\mu}_2, t\}$

the *right* model for $Z = \{z_1, z_2, \ldots z_T\}$?

Problem: In nested models, the model with more parameters does *always* capture the dynamic of the observed time series better!

Approaches:

- Likelihood partition with penalty term for parameters.
- Hypothesis test (Likelihood-ratio statistics).
- (Objective) Bayesian strategies

First: How to parameterise a single linear SDE?

Time series generated by linear SDE's...

$$\dot{z} = F(z - \mu) + \Sigma \dot{W} \quad z \in \mathbf{R}^n$$

$$Z = \{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_T\} \text{ with } \boldsymbol{z}_k = \boldsymbol{z}((k-1)\tau)$$

The solution of a linear SDE is a Gaussian process, iff the initial value is constant or Gaussian distributed (cf. Arnold 1974).

E.g.: $p(\boldsymbol{z}_{t+1} \mid \boldsymbol{z}_t) = |2\pi R(\tau)|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \operatorname{tr}\left((\boldsymbol{z}_{\cdot,t+1} - \boldsymbol{\mu}_t)(\boldsymbol{z}_{\cdot,t+1} - \boldsymbol{\mu}_t)' R(\tau)^{-1}\right)\right)$

$$\boldsymbol{\mu}_t = \boldsymbol{\mu} + \exp(\tau F)(\boldsymbol{z}_t - \boldsymbol{\mu})$$

$$R(\tau) = \int_0^\tau \exp(-F(\tau - s))\Sigma\Sigma' \exp(-F'(\tau - s))ds$$

Allows easy construction of a Likelihood function.

But analytical maximisation wrt. $F, \Sigma, \boldsymbol{\mu}$ is not possible.

Alternative: maximisation wrt. $\theta = (\exp(\tau F), \Sigma\Sigma', \boldsymbol{\mu})$ (Horenko/Schütte 2008)

... are VAR prozcsses.

Separation from trend and random noise shows the autoregressive structure...

$$\begin{aligned} \boldsymbol{z}_{t+1} &= \mathcal{N}(\boldsymbol{\mu} + \exp(\tau F)(\boldsymbol{z}_t - \boldsymbol{\mu}), R) \\ &= (I - \exp(\tau F))\boldsymbol{\mu} + \exp(\tau F)\boldsymbol{z}_t + \mathcal{N}(\mathbf{0}, R) \end{aligned}$$

...enables a compact notation...

$$Y = \Phi X + \epsilon.$$

$$\underset{n+1 \text{ x } T-1}{X} = \begin{pmatrix} 1 & \cdots & 1 \\ \boldsymbol{z}_1 & \cdots & \boldsymbol{z}_{T-1} \end{pmatrix} \qquad \underset{n \text{ x } n+1}{\varPhi} = \begin{pmatrix} (I - \exp(\tau F))\boldsymbol{\mu} & \exp(\tau F) \end{pmatrix}$$

$$\underset{n \text{ x } T-1}{Y} = \begin{pmatrix} \boldsymbol{z}_2 & \cdots & \boldsymbol{z}_T \end{pmatrix} \qquad \underset{n \text{ x } T-1}{\epsilon} = \begin{pmatrix} \mathcal{N}(\mathbf{0}, R) & \cdots & \mathcal{N}(\mathbf{0}, R) \end{pmatrix}$$

...and provides a likelihood function in $\varPhi, R$

$$L(\varPhi, R|Z) = \left(\frac{1}{\sqrt{|2\pi R|}}\right)^{(T-1)} \exp\left(-\frac{1}{2}\operatorname{tr}((Y - \varPhi X)(Y - \varPhi X)' R^{-1})\right)$$

Easy matrix calculus shows:

$$\begin{aligned} \hat{\Phi} &= Y X'(X X')^{-1} \\ \hat{R} &= (Y - \hat{\Phi} X)(Y - \hat{\Phi} X)'/(T - 1) \end{aligned}$$

Up to now we have a VAR(1) model:

$$z_{t+1} = A_0\boldsymbol{\mu} + A_1 z_t + \mathcal{N}(\mathbf{0}, R)$$

$$A_0 := (I - \exp(\tau F))\boldsymbol{\mu}$$

$$A_1 := \exp(\tau F)$$

Generalisation to VAR(p):

$$z_{t+1} = A_0\boldsymbol{\mu} + \sum_{i=1}^{p} A_i z_{t-i+1} + \mathcal{N}(\mathbf{0}, R)$$

Results in a marginal change of estimators

$$X := \begin{pmatrix} 1 & \dots & 1 \\ z_1 & \dots & z_{T-p} \\ \vdots & & \vdots \\ z_p & \dots & z_{T-1} \end{pmatrix}$$

$$Y := (z_{p+1}, \dots, z_T)$$

$$\Phi = \begin{pmatrix} A_0\boldsymbol{\mu} & A_1 & A_2 & \dots & A_p \end{pmatrix}$$

$$\hat{\Phi} = YX'(XX')^{-1}$$

$$\hat{R} = (Y - \hat{\Phi}X)(Y - \hat{\Phi}X)'/(T-p)$$

Allows modelling of non-Markovian effects!

Moment Matrix

Numerically the estimator $\hat{\Phi} = YX'(XX')^{-1}$ is bad.

Better: use the Cholesky factorisation of the (regularised) moment matrix

$$M = M(Z) = \begin{pmatrix} XX' & XY' \\ YX' & YY' \end{pmatrix} = \begin{pmatrix} U'_{11}U_{11} & U'_{11}U_{12} \\ U'_{12}U_{11} & U'_{12}U_{12} + U'_{22}U_{22} \end{pmatrix} = U'U$$

Dimension: $n(p+1)+1$!

$$\hat{\Phi} = (U_{11}^{-1}U_{12})'$$

$$\hat{R} = \frac{1}{T-p} U'_{22}U_{22}$$

(Neumaier/ Schneider 2001)

- M contains all statistical relevant information of the time series.

$$L(\Phi, R|Z) = L(\Phi, R|M) = \left(\frac{1}{\sqrt{|2\pi R|}}\right)^m \cdot \exp\left(-\frac{1}{2}\operatorname{tr}\left((M_{22} - M_{21}\Phi' - \Phi M_{12} + \Phi M_{22}\Phi')R^{-1}\right)\right)$$

- Information from different time series can be combined

$$M(Z_1, Z_2) = M(Z_1) + M(Z_2)$$

H_0 : $Z = \{z_1, z_2, \ldots z_T\}$ was generated by a **single** VAR(p) process.

H_t : the parameters change at time t ($p+1 \leq t \leq T$).

Use Bayesian formula to compute the most probable model:

$$\mathbb{P}[H_t|Z] = \frac{\mathbb{P}[Z|H_t]\,\mathbb{P}[H_t]}{\sum \mathbb{P}[Z|H_j]\,\mathbb{P}[H_j]}$$

$\mathbb{P}[H_t]$: Prior knowledge about model, set e.g.

$$\mathbb{P}[H_0] = \frac{1}{2} \text{ (no change)} \quad \mathbb{P}[H_t] = \frac{1}{2(T-p)} \quad t > 0 \text{ (change)}$$

$\mathbb{P}[Z|H_t]$, obtained via integration in parameter space:

$$\mathbb{P}[Z|H_0] = \int p(Z|\Phi_1, R_1)\pi_1(\Phi_1, R_1)d\Phi_1 dR_1$$

$$\mathbb{P}[Z|H_t] = \int p(Z_1|\Phi_1, R_1)\pi_1(\Phi_1, R_1)p(Z_2|\Phi_2, R_2)\pi_2(\Phi_2, R_2)d\Phi_1 dR_1 d\Phi_2 dR_2$$

Problem: specification of the prior distributions $\pi_1(\Phi_1, R_1)$ and $\pi_2(\Phi_2, R_2)$?

Standard priors are only defined up to a constant, e.g.

$$\pi_1(\Phi, R) = c_1 |R|^{-\frac{d+1}{2}}$$
$$\pi_2(\Phi, R) = c_2 |R|^{-\frac{d+1}{2}}$$

$$\longrightarrow \quad \mathbb{P}[H_t|Z] = c \frac{\mathbb{P}[Z|H_t]\,\mathbb{P}[H_i]}{\sum_j \mathbb{P}[Z|H_j]\,\mathbb{P}[H_j]} \quad ?$$

Exclude H_0 to avoid the model selection problem:

$$\mathbb{P}[H_t|Z] = \frac{\cancel{c_1}\,\mathbb{P}[Z|H_t]\,\cancel{\mathbb{P}[H_t]}}{\cancel{c_1}\sum_{j\geq 1}\mathbb{P}[Z|H_j]\,\cancel{\mathbb{P}[H_j]}}$$

Evaluation is analytical possible in terms moment matrices

$$M_1 = M(Z_1) \qquad Z_1 = \{z_1, \ldots, z_t\}$$
$$M_2 = M(Z_2) \qquad Z_2 = \{z_{t+1}, \ldots, z_T\}$$

$$I[M] = \pi^{\frac{d(d-1)}{2}} |U_{11}|^{-d} |\sqrt{\pi}\, U_{22}|^{dp+1-m} \prod_{j=1}^{d} \Gamma\left(\frac{m-dp-j}{2}\right)$$

The most likely change point t_* can be obtained easily.

Still needed: a decision between H_0 and H_{t_*}.

Fractional Bayes: Take a fraction of the likelihood to normalise the prior, i.e.

$$\pi(\Phi, R) \propto L^b(\Phi, R|Z).|R|^{-\frac{d+1}{2}} \quad 0 \leq b \leq 1$$

$$L(\Phi, R|Z) \rightarrow L^{1-b}(\Phi, R|Z)$$

Again an analytical solution: $\mathbb{P}[H_{t_*}] = \dfrac{I[M_1]I[M_2]}{I[M_1]I[M_2] + I[M_1 + (1-b)M_2]}$

An online change point algorithm can be constructed:

$Get\ data\ Z_I = \{\boldsymbol{z_0}, \ldots, \boldsymbol{z}_{t_I}\}$.
$\pi_1(\Phi, R) \propto L(\Phi, R|Z_I)$
$t_A \leftarrow t_I + 1$
$t_E \leftarrow t_I + t_w$
$\mathbb{P}[\text{change}] \leftarrow 0$

while $\mathbb{P}[change] < \alpha$ **do**
- $Get\ new\ data\ Z_W = \{\boldsymbol{z}_{t_A}, \ldots, \boldsymbol{z}_{t_E}\}$.
- $Determine\ candidate\ change\ point\ t_*$
- $Z_1 = \{\boldsymbol{z}_A, \ldots, \boldsymbol{z}_{t_*-1}\}$
- $Z_2 = \{\boldsymbol{z}_{t_*}, \ldots, \boldsymbol{z}_{t_E}\}$
- $\pi_2(\Phi, R) \propto L^b(\Phi, R|Z_2)$
- $L(\Phi, R|Z_2) \leftarrow L^{1-b}(\Phi, R|Z_2)$
- $Compute\ \mathbb{P}[change]\ according\ to\ Bayes.$
- $t_E \leftarrow t_E + t_w$