Filter Based Volatility in the HMM

Elisabeth Leoff   Jörn Sass

TU Kaiserslautern

May 26 2014

Overview

Technische Universität Kaiserslautern

Hidden Markov Model in Continuous Time

- Observation process $Y = (Y_t)_{t\in[0,T]}$ on $(\Omega, \mathcal{A}, \mathbb{P})$ with

$$Y_t = \int_0^t \mu_s \mathrm{d}s + W_t.$$

Hidden Markov Model in Continuous Time

- Observation process $Y = (Y_t)_{t \in [0,T]}$ on $(\Omega, \mathcal{A}, \mathbb{P})$ with

$$Y_t = \int_0^t \mu_s \mathrm{d}s + W_t.$$

- Y_t models a return process.

Hidden Markov Model in Continuous Time

- Observation process $Y = (Y_t)_{t \in [0,T]}$ on $(\Omega, \mathcal{A}, \mathbb{P})$ with

$$Y_t = \int_0^t \mu_s \mathrm{d}s + W_t.$$

- Y_t models a return process.
- Drift μ_t is governed by signal X_t, a Markov jump process with state space $\{e_1, \ldots, e_d\}$, so $\mu_t = \langle \mu, X_t \rangle$, and rate matrix Q.

Technische Universität Kaiserslautern

Hidden Markov Model in Continuous Time

- Observation process $Y = (Y_t)_{t \in [0,T]}$ on $(\Omega, \mathcal{A}, \mathbb{P})$ with

$$Y_t = \int_0^t \mu_s \mathrm{d}s + W_t.$$

- Y_t models a return process.
- Drift μ_t is governed by signal X_t, a Markov jump process with state space $\{e_1, \ldots, e_d\}$, so $\mu_t = \langle \mu, X_t \rangle$, and rate matrix Q.
- The volatility is constant.

Technische Universität Kaiserslautern

Hidden Markov Model in Continuous Time

- Observation process $Y = (Y_t)_{t\in[0,T]}$ on $(\Omega, \mathcal{A}, \mathbb{P})$ with

$$Y_t = \int_0^t \mu_s \mathrm{d}s + W_t.$$

- Y_t models a return process.
- Drift μ_t is governed by signal X_t, a Markov jump process with state space $\{e_1, \dots, e_d\}$, so $\mu_t = \langle \mu, X_t \rangle$, and rate matrix Q.
- The volatility is constant.
- W Brownian motion, independent of signal X.

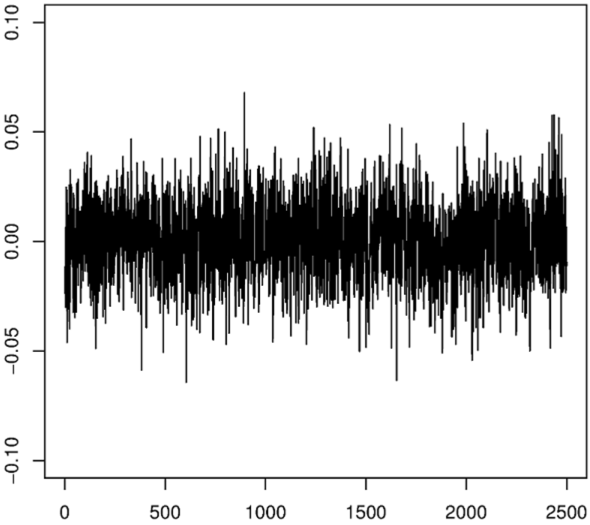
TECHNISCHE UNIVERSITÄT KAISERSLAUTERN

Hidden Markov Model in Continuous Time

- Observation process $Y = (Y_t)_{t\in[0,T]}$ on $(\Omega, \mathcal{A}, \mathbb{P})$ with

$$Y_t = \int_0^t \mu_s \mathrm{d}s + W_t.$$

- Y_t models a return process.
- Drift μ_t is governed by signal X_t, a Markov jump process with state space $\{e_1, \dots, e_d\}$, so $\mu_t = \langle \mu, X_t \rangle$, and rate matrix Q.
- The volatility is constant.
- W Brownian motion, independent of signal X.
- We are interested in the filter $\mathbb{E}\left[X_t \mid \mathcal{Y}_t\right]$, where $\mathcal{Y}_t = \sigma(Y_s | s \leq t)$ is the observable information.


0.10
0.05
0.00
−0.05
−0.10
0
500
1000
1500
2000
2500

Change of Measure

- We introduce a change of measure using the density

$$\left.\frac{\mathrm{d}\tilde{\mathbb{P}}}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{F}_t} = Z_t = \exp\left(\int_0^t \langle \mu, X_s \rangle \mathrm{d}W_s - \frac{1}{2}\int_0^t \langle \mu, X_s \rangle^2 \mathrm{d}s\right).$$

Change of Measure

- We introduce a change of measure using the density

$$\left.\frac{\mathrm{d}\tilde{\mathbb{P}}}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{F}_t} = Z_t = \exp\left(\int_0^t \langle \mu, X_s\rangle \mathrm{d}W_s - \frac{1}{2}\int_0^t \langle \mu, X_s\rangle^2 \mathrm{d}s\right).$$

- Define the unnormalized filter as

$$\rho_t(X) := \tilde{\mathbb{E}}\left[Z_t^{-1} X_t | \mathcal{Y}_t\right] \propto \mathbb{E}\left[X_t \mid \mathcal{Y}_t\right].$$

Filtering Equations

- For the unnormalized filter we have that

$$\mathrm{d}\rho_t(X) = Q^T \rho_t(X)\mathrm{d}t + B\rho_t(X)\mathrm{d}Y_t,$$

where $B = \mathrm{Diag}(\mu)$.

Filtering Equations

- For the unnormalized filter we have that

$$\mathrm{d}\rho_t(X) = Q^T \rho_t(X)\mathrm{d}t + B\rho_t(X)\mathrm{d}Y_t,$$

where $B = \text{Diag}(\mu)$.

- In general, filtering equations are not finite dimensional.

Portfolio Optimization

- Portfolio optimization is possible in the HMM (Haussmann/Sass, 2004).

Portfolio Optimization

- Portfolio optimization is possible in the HMM (Haussmann/Sass, 2004).
- Goal is to maximize $\mathbb{E}[U(V_T^\pi)]$ for wealth process V^π and utility U.

Portfolio Optimization

- Portfolio optimization is possible in the HMM (Haussmann/Sass, 2004).
- Goal is to maximize $\mathbb{E}[U(V_T^\pi)]$ for wealth process V^π and utility U.
- For $U = \log$ this leads to the optimal risky fraction

$$\pi_t = \frac{\langle \mu, \mathbb{E}[X_t|\mathcal{Y}_t] \rangle}{\sigma^2}.$$

Portfolio Optimization

- Portfolio optimization is possible in the HMM (Haussmann/Sass, 2004).
- Goal is to maximize $\mathbb{E}[U(V_T^\pi)]$ for wealth process V^π and utility U.
- For $U = \log$ this leads to the optimal risky fraction

$$\pi_t = \frac{\langle \mu, \mathbb{E}[X_t | \mathcal{Y}_t] \rangle}{\sigma^2}.$$

- The optimal strategy depends on the filter.

Properties of the HMM

- Filter equations are finite dimensional, so the filters can be computed from finitely many SDEs.

Properties of the HMM

- Filter equations are finite dimensional, so the filters can be computed from finitely many SDEs.
- Parameters can be estimated using the EM-algorithm (Elliott, 1992).

Properties of the HMM

- Filter equations are finite dimensional, so the filters can be computed from finitely many SDEs.
- Parameters can be estimated using the EM-algorithm (Elliott, 1992).
- Robustifications are available which eliminate the stochastic integral from the filtering equation (James/Krishnamurthy/Le Gland, 1992).

Properties of the HMM

- Filter equations are finite dimensional, so the filters can be computed from finitely many SDEs.
- Parameters can be estimated using the EM-algorithm (Elliott, 1992).
- Robustifications are available which eliminate the stochastic integral from the filtering equation (James/Krishnamurthy/Le Gland, 1992).
- Filters can be derived for any $\sigma > 0$.

Properties of the HMM

- Filter equations are finite dimensional, so the filters can be computed from finitely many SDEs.
- Parameters can be estimated using the EM-algorithm (Elliott, 1992).
- Robustifications are available which eliminate the stochastic integral from the filtering equation (James/Krishnamurthy/Le Gland, 1992).
- Filters can be derived for any $\sigma > 0$.
- But: σ has to be constant.

Properties of the HMM

- Filter equations are finite dimensional, so the filters can be computed from finitely many SDEs.
- Parameters can be estimated using the EM-algorithm (Elliott, 1992).
- Robustifications are available which eliminate the stochastic integral from the filtering equation (James/Krishnamurthy/Le Gland, 1992).
- Filters can be derived for any $\sigma > 0$.
- But: σ has to be constant.
- So 'Stylized Facts' as e.g jumps in the volatility or volatility clustering cannot be modeled using a HMM!

Markov Switching Model

- Drift **and** volatility are governed by Markov chain X.

Markov Switching Model

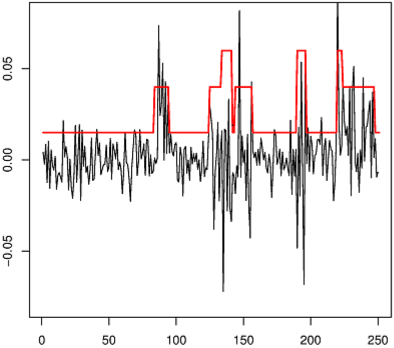
- Drift **and** volatility are governed by Markov chain X.
- Observations $Y = (Y_t)_{t\in[0,T]}$ are given by

$$Y_t = \int_0^t \langle \mu, X_s \rangle \mathrm{d}s + \int_0^t \langle \sigma, X_s \rangle \mathrm{d}W_s.$$

Markov Switching Model

- Drift **and** volatility are governed by Markov chain X.
- Observations $Y = (Y_t)_{t\in[0,T]}$ are given by

$$Y_t = \int_0^t \langle \mu, X_s\rangle \mathrm{d}s + \int_0^t \langle \sigma, X_s\rangle \mathrm{d}W_s.$$



Properties of the MSM

- Many 'Stylized Facts' can be modeled.

Properties of the MSM

- Many 'Stylized Facts' can be modeled.
- But: In theory X_t can be observed via $[Y]_t$.

Properties of the MSM

- Many 'Stylized Facts' can be modeled.
- But: In theory X_t can be observed via $[Y]_t$.
- There is no filtering problem in the continuous time MSM!

Properties of the MSM

- Many 'Stylized Facts' can be modeled.
- But: In theory X_t can be observed via $[Y]_t$.
- There is no filtering problem in the continuous time MSM!
- The optimal risky fraction for log utility is (Bäuerle/Rieder, 2004)

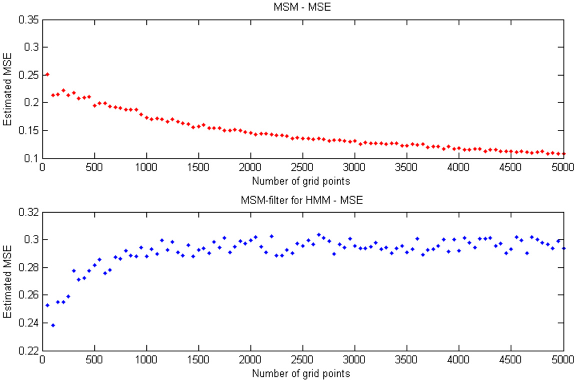
$$\pi_t = \frac{\langle \mu, X_t \rangle}{\langle \sigma, X_t \rangle^2}.$$

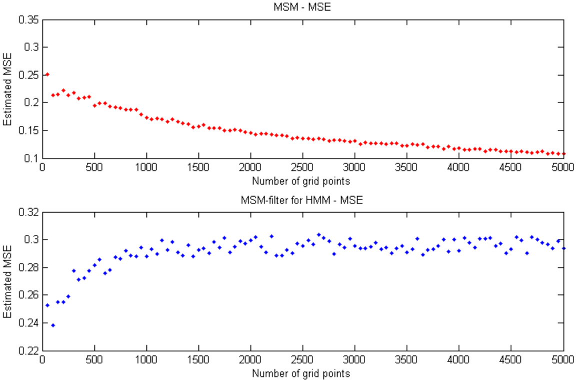
Properties of the MSM

- Many 'Stylized Facts' can be modeled.
- But: In theory X_t can be observed via $[Y]_t$.
- There is no filtering problem in the continuous time MSM!
- The optimal risky fraction for log utility is (Bäuerle/Rieder, 2004)

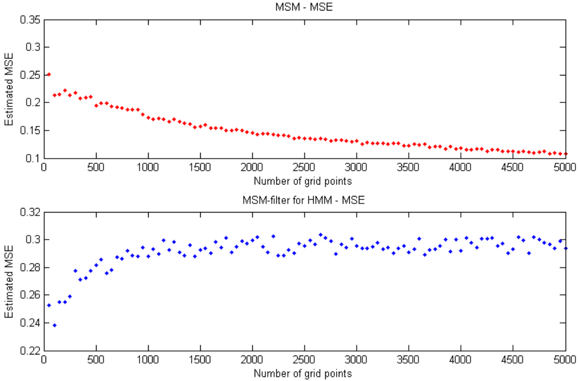
$$\pi_t = \frac{\langle \mu, X_t \rangle}{\langle \sigma, X_t \rangle^2}.$$

- For finer discretizations, the MSE in the MSM behaves differently than in the HMM.


MSM - MSE
Estimated MSE
0.35
0.3
0.25
0.2
0.15
0.1
0
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
Number of grid points
MSM-filter for HMM - MSE
Estimated MSE
0.32
0.3
0.28
0.26
0.24
0.22
Number of grid points



- Idea: Different dynamics for σ.



- Idea: Different dynamics for σ.
- Model volatility as depending on an observable process.

Filter Based Volatility Model

- Observations Y are

$$Y_t = \int_0^t \langle \mu, X_s \rangle \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s.$$

Filter Based Volatility Model

- Observations Y are

$$Y_t = \int_0^t \langle \mu, X_s \rangle \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s.$$

- Change of measure with density

$$\mathrm{d}Z_t = -Z_t \sigma_t^{-1} \langle \mu, X_t \rangle \mathrm{d}W_t$$

Filter Based Volatility Model

- Observations Y are

$$Y_t = \int_0^t \langle \mu, X_s \rangle \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s.$$

- Change of measure with density

$$\mathrm{d}Z_t = -Z_t \sigma_t^{-1} \langle \mu, X_t \rangle \mathrm{d}W_t$$

leads to

$$Y_t = \int_0^t \sigma_s \mathrm{d}\tilde{W}_s$$

Filter Based Volatility Model

- Observations Y are

$$Y_t = \int_0^t \langle \mu, X_s \rangle \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s.$$

- Change of measure with density

$$\mathrm{d}Z_t = -Z_t \sigma_t^{-1} \langle \mu, X_t \rangle \mathrm{d}W_t$$

leads to

$$Y_t = \int_0^t \sigma_s \mathrm{d}\tilde{W}_s$$

where $\mathrm{d}\tilde{W}_t = \mathrm{d}W_t + \sigma_t^{-1} \langle \mu, X_t \rangle \mathrm{d}t$ $\tilde{\mathbb{P}}$-Brownian motion.

Filter Based Volatility

- Consider a model for $\sigma_t = \sigma(\xi_t)$ with

$$\mathrm{d}\xi_t = \nu(\xi_t)\mathrm{d}t + \tau(\xi_t)\mathrm{d}\tilde{W}_t.$$

Filter Based Volatility

- Consider a model for $\sigma_t = \sigma(\xi_t)$ with

$$\mathrm{d}\xi_t = \nu(\xi_t)\mathrm{d}t + \tau(\xi_t)\mathrm{d}\tilde{W}_t.$$

- The equation for the unnormalized filter is

$$\mathrm{d}\rho_t(X) = Q^T \rho_t(X)\mathrm{d}t + \frac{1}{\sigma_t^2} B\rho_t(X)\mathrm{d}Y_t.$$

Filter Based Volatility

- Consider a model for $\sigma_t = \sigma(\xi_t)$ with

$$\mathrm{d}\xi_t = \nu(\xi_t)\mathrm{d}t + \tau(\xi_t)\mathrm{d}\tilde{W}_t.$$

- The equation for the unnormalized filter is

$$\mathrm{d}\rho_t(X) = Q^T\rho_t(X)\mathrm{d}t + \frac{1}{\sigma_t^2}B\rho_t(X)\mathrm{d}Y_t.$$

- Since $\mathrm{d}Y_t = \sigma_t\mathrm{d}\tilde{W}_t$,

Filter Based Volatility

- Consider a model for $\sigma_t = \sigma(\xi_t)$ with

$$\mathrm{d}\xi_t = \nu(\xi_t)\mathrm{d}t + \tau(\xi_t)\mathrm{d}\tilde{W}_t.$$

- The equation for the unnormalized filter is

$$\mathrm{d}\rho_t(X) = Q^T \rho_t(X)\mathrm{d}t + \frac{1}{\sigma_t^2} B\rho_t(X)\mathrm{d}Y_t.$$

- Since $\mathrm{d}Y_t = \sigma_t \mathrm{d}\tilde{W}_t$, we can choose

$$\xi_t = \rho_t(X).$$

Approximating the MSM

- Assume returns in both models are depending on the same Markov chain X and Brownian motion W.

Approximating the MSM

- Assume returns in both models are depending on the same Markov chain X and Brownian motion W.
- Difference of the returns is

$$Y_t^{MSM} - Y_t^{FV} = \int_0^t \left(\langle \sigma, X_t \rangle - \sigma(\xi_t)\right) \mathrm{d}W_s.$$

Approximating the MSM

- Assume returns in both models are depending on the same Markov chain X and Brownian motion W.
- Difference of the returns is

$$Y_t^{MSM} - Y_t^{FV} = \int_0^t \left(\langle \sigma, X_t \rangle - \sigma(\xi_t) \right) \mathrm{d}W_s.$$

- Want to minimize the (squared) distance over all $\mathcal{Y}^{FV}$-adapted processes $\tilde{\xi}$:

Approximating the MSM

- Assume returns in both models are depending on the same Markov chain X and Brownian motion W.
- Difference of the returns is

$$Y_t^{MSM} - Y_t^{FV} = \int_0^t \left(\langle \sigma, X_t \rangle - \sigma(\xi_t) \right) \mathrm{d}W_s.$$

- Want to minimize the (squared) distance over all $\mathcal{Y}^{FV}$-adapted processes $\tilde{\xi}$:

$$\mathbb{E}[(Y_t^{MSM} - Y_t^{FV})^2] \stackrel{!}{=} \min_{\tilde{\xi}} \mathbb{E}\left[\left(\int_0^t \langle \sigma, X_t \rangle - \sigma(\tilde{\xi}_t) \mathrm{d}W_s \right)^2 \right].$$

Approximating the MSM

- This results in

$$\sigma(\xi_t) = \mathbb{E}[\langle \sigma, X_t \rangle | \mathcal{Y}_t^{FV}]$$

Approximating the MSM

- This results in

$$\sigma(\xi_t) = \mathbb{E}[\langle \sigma, X_t \rangle | \mathcal{Y}_t^{FV}] = \langle \sigma, \mathbb{E}[X_t | \mathcal{Y}_t^{FV}] \rangle.$$

Approximating the MSM

- This results in

$$\sigma(\xi_t) = \mathbb{E}[\langle \sigma, X_t \rangle | \mathcal{Y}_t^{FV}] = \langle \sigma, \mathbb{E}[X_t | \mathcal{Y}_t^{FV}] \rangle.$$

- So: The general ansatz of observable volatility approximating the MSM leads to a function of the filter!

Approximating the MSM

- This results in

$$\sigma(\xi_t) = \mathbb{E}[\langle \sigma, X_t \rangle | \mathcal{Y}_t^{FV}] = \langle \sigma, \mathbb{E}[X_t | \mathcal{Y}_t^{FV}] \rangle.$$

- So: The general ansatz of observable volatility approximating the MSM leads to a function of the filter!
- This function corresponds exactly to the function of X in the MSM.

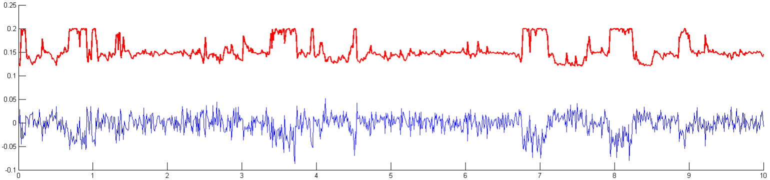
Approximating the MSM

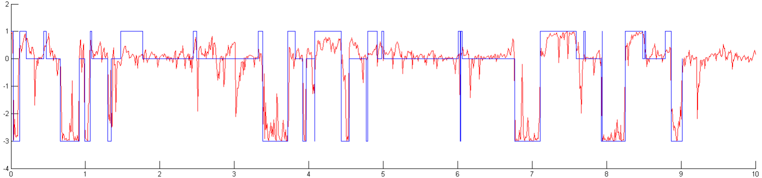
- This results in

$$\sigma(\xi_t) = \mathbb{E}[\langle \sigma, X_t \rangle | \mathcal{Y}_t^{FV}] = \langle \sigma, \mathbb{E}[X_t | \mathcal{Y}_t^{FV}] \rangle.$$

- So: The general ansatz of observable volatility approximating the MSM leads to a function of the filter!
- This function corresponds exactly to the function of X in the MSM.
- In particular, it is linear in the filter.

Linear Model





Properties of the Model

- Model captures the same 'Stylized Facts' as the MSM in general.

Properties of the Model

- Model captures the same 'Stylized Facts' as the MSM in general.
- But some are less pronounced, e.g. Heavy Tails.

Properties of the Model

- Model captures the same 'Stylized Facts' as the MSM in general.
- But some are less pronounced, e.g. Heavy Tails.
- Portfolio optimization can be done for several utility functions (Sass/Haussmann,2004).

Properties of the Model

- Model captures the same 'Stylized Facts' as the MSM in general.
- But some are less pronounced, e.g. Heavy Tails.
- Portfolio optimization can be done for several utility functions (Sass/Haussmann,2004).
- Robust discretizations are available.

Summary

The filter based volatility model

- ...has non-constant volatility.

Summary

The filter based volatility model

- ...has non-constant volatility.
- ...has a finite dimensional filtering equation.

Summary

The filter based volatility model

- ...has non-constant volatility.
- ...has a finite dimensional filtering equation.
- ...approximates the MSM.

Summary

The filter based volatility model

- ...has non-constant volatility.
- ...has a finite dimensional filtering equation.
- ...approximates the MSM.
- ...allows for portfolio optimization and parameter estimation with the EM-algorithm.

Summary

The filter based volatility model

- ...has non-constant volatility.
- ...has a finite dimensional filtering equation.
- ...approximates the MSM.
- ...allows for portfolio optimization and parameter estimation with the EM-algorithm.

But: Some 'Stylized Facts' are less pronounced.

Literature

N. Bäuerle and U. Rieder.
Portfolio optimization with Markov-modulated stock prices and interest rates.
IEEE Transactions on Automatic control, 49:442–447, 2004.

R. J. Elliott.
New Finite-Dimensional Filters and Smoothers for Noisily Observed Markov Chains.
IEEE Trans. Inf. Theor., 39(1):265–271, 1993.

M. R. James, V. Krishnamurthy, and F. Le Gland.
Time Discretization of Continuous-Time Filters and Smoothers for HMM Parameter Estimation.
IEEE Transactions on Information Theory, 42:593–605, 1996.

J. Sass and U. G. Haussmann.
Portfolio optimization under partial information: Stochastic volatility in a hidden Markov model.
In R. Fahrion D. Ahr, editor, *Operations Research Proceedings 2003*, pages 387–394. Springer, Berlin, 2004.

J. Sass and U.G. Haussmann.
Optimizing the terminal wealth under partial information: The drift process as a continuous time markov chain.
Finance and Stochastics, 8:553–577, 2004.