Continuous-Time Regime Switching Models, Portfolio Optimization and Filter-Based Volatility

Jörn Sass

joint work with . . . , Elisabeth Leoff, Vikram Krishnamurthy

sass@mathematik.uni-kl.de
University of Kaiserslautern, Germany

Wien, March 6, 2015

Regime switching, portfolio optimization, filter-based volatility

- Markov switching and hidden Markov models (MSMs and HMMs)
- Partial information and filtering
- Portfolio optimization
- Continuous versus discrete time
- HMMs with non-constant volatility

- Markov switching and hidden Markov models (MSMs and HMMs)

A continuous-time Markov switching model (MSM)

- Observation process $R = (R_t)_{t\in[0,T]}$, e.g. stock returns,

$$R_t = \int_0^t \mu_s \, ds + \int_0^t \sigma_s \, dW_s$$

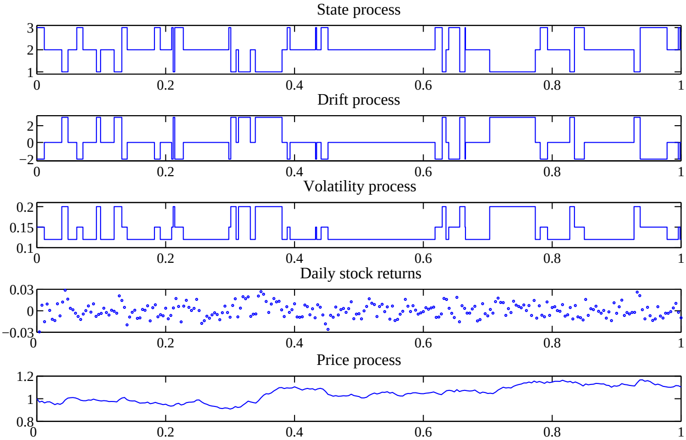
- Drift $\mu_t = b^\top Y_t = \sum b_i Y_t^i$, $b \in \mathbb{R}^d$, and volatility $\sigma_t = a^\top Y_t$, $a \in \mathbb{R}^d_{>0}$
- $Y = (Y_t)_{t\in[0,T)}$ continuous-time Markov chain with states $\{e_1, \dots, e_d\}$
- W standard Brownian motion, independent of Y
- Jumps are governed by rate matrix $Q \in \mathbb{R}^{d\times d}$
  - Diagonal: Exponential rate of leaving state e_k,

$$\lambda_k = -Q_{kk} = \sum_{l\neq k} Q_{kl} < \infty$$

  - Conditional transition probability:

$$\mathrm{P}(Y_t = e_l \mid Y_{t-} = e_k, Y_t \neq Y_{t-}) = Q_{kl}/\lambda_k$$

Example: Simulated data



$$\Delta t = \frac{1}{250},\ b^\top = (3, 0, -2),\ a^\top = (0.20, 0.12, 0.15),\ Q = \begin{pmatrix} -70 & 40 & 30 \\ 20 & -40 & 20 \\ 30 & 50 & -80 \end{pmatrix}$$

Exmple: Daily returns of stock indices

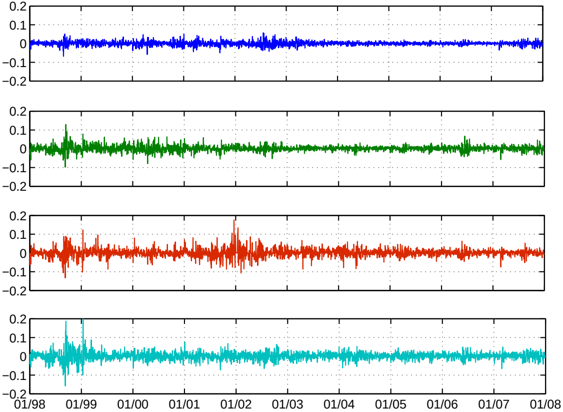
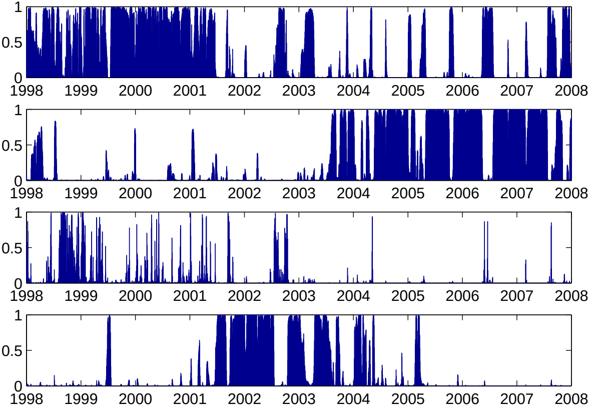


Figure: Daily returns over 10 years for S&P 500, IPC, MerVal, Bovespa

Estimation of state probabilities



State probabilities for states 1 to 4

Estimation by MCMC methods in Hahn/Frühwirth-Schnatter/S. (2010)

Properties, motivation of MSM, HMM

- Properties, see Rydén/Teräsvirta/Åsbrink (1998), Timmermann (2000):
  - Wide ranges for skewness, kurtosis, tails; leverage and volatility clustering
  - Negative: No jumps, decay of autocorrelation of $|\Delta R|$, ΔR^2 too fast
- Interpretation:
  - State process models unobservable underlying economic variable
  - Rare jumps – structural breaks, frequent jumps – arrival of news
- Many applications, e.g. in biophysics, finance, signal processing
- MSM and HMM: Since

$$[R]_t = \int_0^t \sigma_s^2 ds = \sum_{i=1}^d a_i^2 \int_0^t \mathbf{1}_{\{Y_s = e_i\}} ds,$$

  we distinguish
  - MSM if $a_i \neq a_j$ for all i, j.
  - HMM if $a_1 = \ldots = a_d$ (hidden Markov model).

- Partial information and filtering

Partial information

- HMM is MSM with $a_1 = \ldots = a_d = \sigma$. In the HMM we observe

$$R_t = \int_0^t \mu_s \, ds + \sigma \, W_t, \quad \text{where} \quad \mu_s = b^\top Y_s.$$

- An investor observing R has partial information only, information at t is

$$\mathcal{F}_t^R \subsetneq \mathcal{F}_t.$$

- Then, the best estimator for μ_t is the filter

$$\hat{\mu}_t = \mathrm{E}[\mu_t \,|\, \mathcal{F}_t^R] = b^\top \mathrm{E}[Y_t \,|\, \mathcal{F}_t^R] = b^\top \hat{Y}_t,$$

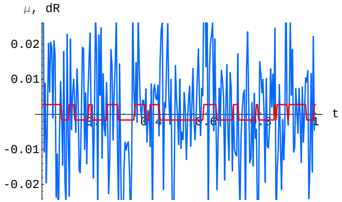
  where $\hat{Y}_t = \mathrm{E}[Y_t \,|\, \mathcal{F}_t^R]$ is the Wonham filter for Y_t.
- In the MSM with switching volatility $\sigma_t = a^\top Y_t$, Y_t can in theory be observed via $[R]_t$. Thus there is no filtering problem in the MSM, Y is not hidden!
- For time-discrete observations Y is hidden for both constant and switching σ.

Filtering in the HMM

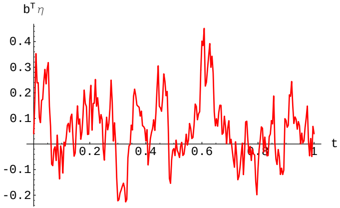
- We consider $dR_t = \mu_t\, dt + \sigma\, dW_t$ and use $dZ_t = -Z_t(\sigma^{-1}\mu_t)^\top\, dW_t$.
- Under $\widetilde{P} \sim P$ by $\dfrac{d\widetilde{P}}{dP} = Z_T$, $\widetilde{W} = \sigma^{-1} R$ is Brownian motion indep. of Y.
- We need $\hat{\mu}_t = b^\top \hat{Y}_t$ for $\hat{Y}_t = \mathrm{E}[Y_t \,|\, \mathcal{F}_t^R]$. Let $\hat{Z}_t = \mathrm{E}[Z_t \,|\, \mathcal{F}_t^R]$.
- The unnormalized filter $\rho_t(Y) := \widetilde{\mathrm{E}}[Z_t^{-1} Y_t \,|\, \mathcal{F}_t^R]$ satisfies Zakai-equation

$$d\rho_t(Y) = Q^\top \rho_t(Y)\, dt + \mathrm{Diag}(\rho_t(Y))\, b\, \sigma^{-2} dR_t, \quad \rho_0(Y) = \mathrm{E}[Y_0].$$

- Using $\hat{Z}_t^{-1} = \mathbf{1}^\top \rho_t(Y)$, Bayes' formula yields $\hat{Y}_t = \frac{\rho_t(Y)}{\mathbf{1}^\top \rho_t(Y)}$.



$\mu\, \Delta t$ and daily returns ΔR



Filter $\hat{\mu}$

- Portfolio optimization

Trading in a HMM

- One money market with interest rate 0 and one stock with returns

$$dR_t = \mu_t \, dt + \sigma \, dW_t$$

- X_t wealth (portfolio value) at t.
- $\pi = (\pi_t)_{t \in [0,T]}$ trading strategy
  π_t is fraction of wealth X_t invested in stock.
  π has to be $\mathcal{F}^R$-adapted.
- $X_t = X_t^\pi$ is controlled by π.
- For initial capital $x_0 > 0$ we have

$$dX_t = X_t \, \pi_t dR_t, \quad X(0) = x_0.$$

- $X_t(1 - \pi_t)$ is invested in the money market (self-financing).

Utility maximization

- Evaluation of terminal wealth by increasing, concave utility function U, e.g.

$$U_\alpha(x) = \frac{x^\alpha}{\alpha},\ \alpha < 1, \alpha \neq 1 \quad \text{or} \quad U_0(x) = \log(x).$$

- **Stochastic control problem:** Maximize expected utility

$$\mathrm{E}[U(X_T^\pi)] \quad \text{over admissible} \quad \pi \quad \text{for} \quad x_0 > 0.$$

- For constant μ

$$\pi_t^* = \frac{1}{1-\alpha}\,\frac{\mu}{\sigma^2}, \quad t \in [0, T], \quad \text{Merton strategy}.$$

- For non-constant μ we expect a dependency on $\hat{\mu}_t$ and its dynamics.
- In general $X_T^* = (U')^{-1}(y\hat{Z}_T)$, where $\hat{Z}_T = \mathrm{E}[Z_T \,|\, \mathcal{F}_T^R]$, $\widetilde{\mathrm{E}}[X_T^*] = x_0$.
- π^* from $\int_0^T (\pi_t^*)\sigma d\widetilde{W}_t = X_T^* - x_0 = \int_0^T \mathrm{E}[D_t X_T^* | \mathcal{F}_t^R] d\widetilde{W}_t$ if latter exists.

Optimal trading strategies

- In the HMM (S./Haussmann 2004)

$$\pi_t^* = \frac{1}{(1-\alpha)\mathrm{E}\left[\hat{Z}_{t,T}^{\frac{\alpha}{\alpha-1}} \mid \rho_t\right]} \Big\{ \sigma^{-2} b^\top \hat{Y}_t \mathrm{E}\left[\hat{Z}_{t,T}^{\frac{2\alpha-1}{\alpha-1}} \mid \rho_t\right] + \sigma^{-1}\mathrm{E}\Big[\hat{Z}_{t,T}^{\frac{2\alpha-1}{\alpha-1}} \int_t^T (D_t\rho_{t,s}) b\sigma^{-2} dR_s \,\Big|\, \rho_t\Big]\Big\}.$$

  For $U = \log$ this becomes $\pi_t^* = \sigma^{-2}\hat{\mu}_t = \sigma^{-2} b^\top \hat{Y}_t$.

- In the MSM (Bäuerle/Rieder 2004) for

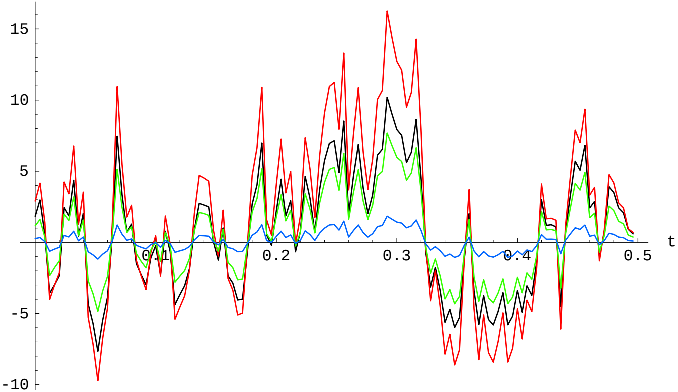
$$\pi_t^* = \frac{1}{1-\alpha} \frac{b^\top Y_t}{(a^\top Y_t)^2}.$$

  For $U = \log$ this becomes $\pi_t^* = \sigma_t^{-2}\mu_t = (a^\top Y_t)^{-2} b^\top Y_t$.

- Continuous versus discrete time

Optimal risky fractions in the HMM

For utility functions $U_0(x) = \log(x)$ and $U_\alpha(x) = x^\alpha/\alpha$, $\alpha < 1$, $\alpha \neq 0$:



Optimal risky fractions π_t^* for $\alpha = 0.2$, log, $\alpha = -0.5$, $\alpha = -5$.

Implementation of optimal strategies

- For maximizing $E[\log(X_T^\pi)]$, the optimal risky fraction is $\pi_t^* = \sigma^{-2}\,\hat{\mu}_t$.
- Constrained strategy: No short selling, no borrowing: Cut off π^* at 0, 1.
- Average log-utilities (500 simulations) for different trading frequencies:

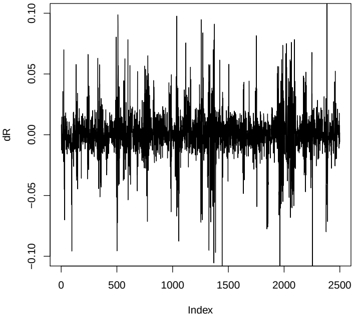
strategy	10/day	5/day	4/day	2/day	daily	every 2 days
constrained	0.261	0.256	0.246	0.230	0.192	0.165

for $d = 2$, $\sigma = 0.4$, $b^\top = (2.5, -1.5)$, $Q_{12} = 60$, $Q_{21} = 40$, i.e. $\mathrm{E}[\mu_t] = 0.1$.

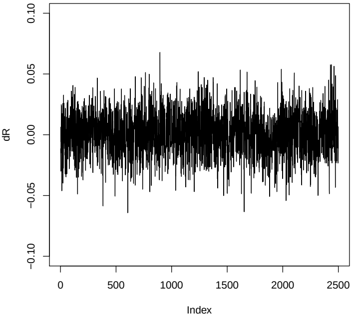
- In discretized model same results as for constrained strategy.
- Thus in the HMM, the discretized model is well approximated by the continuous time model with constraints (or with mild parameters).
- Optimal constrained strategy in continuous-time MSM leads to optimal expected utilities about 0.968 versus 0.192.
- Thus, continuous-time MSM is poor approximation for discrete-time MSM.

Reminder

Reminder: From the econometric properties, the continuous-time MSM is preferable to the continuous-time HMM.



MSM over 10 years



HMM over 10 years

- HMMs with non-constant volatility

MSM versus HMM with non-constant volatility

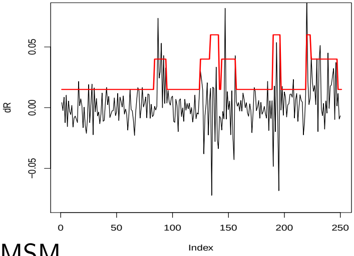
- The continuous-time MSM is a poor approximation for the discrete time model in view of portfolio optimization.
- Idea: Consider a HMM with a non-constant volatility model,

$$dR_t = b^\top Y_t + \sigma_t \, dW_t,$$

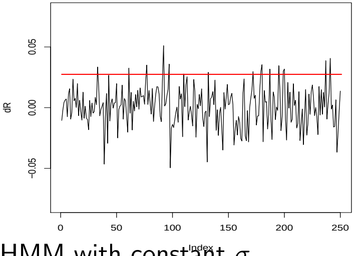
where $\sigma_t = f(\hat{Y}_t)$, as approximation for the MSM.

- This yields consistent continuous-time approximations, since
  - For non-constant σ_t filters can be computed (Haussmann/S. 2004).
  - For non-constant σ_t, optimal strategy π_t^* can be computed as above. It then has an additional term due to the dynamics of σ_t.
  - The dependency can be modelled such that $f(Y_t) = a^\top Y_t$.
- Any dynamic volatility model w.r.t. $\widetilde{W}$ can be used.

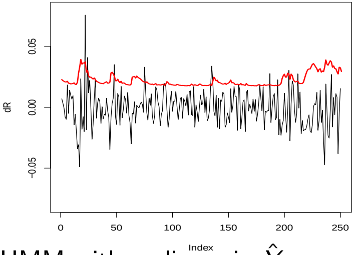
Daily returns and volatility process



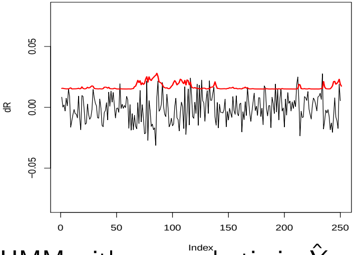
MSM



HMM with constant σ



HMM with σ_t linear in $\hat{Y}_t$



HMM with σ_t quadratic in $\hat{Y}_t$

HMM with non-constant volatility closest to MSM

- Consider for $\mathcal{F}^R$ adapted $(\sigma_t)_{t\in[0,T]}$

$$dR_t = b^\top Y_t dt + a^\top Y_t dW_t \qquad \text{and} \qquad dR_t^H = b^\top Y_t dt + \sigma_t dW_t$$

- The mean squared distance of the return processes is

$$\mathrm{MSE}(R, R^H) = \frac{1}{T}\,\mathrm{E}\left[\int_0^T (R_t - R_t^H)^2 dt\right].$$

- We have

$$\mathrm{MSE(R, R^H)} = \frac{1}{T}\int_0^T \int_0^t \mathrm{E}\left[(a^\top Y_s - \sigma_s)^2\right] ds\, dt.$$

- This is minimized by

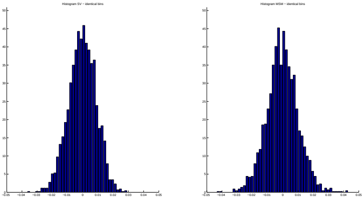
$$\sigma_t = \mathrm{E}\left[a^\top Y_t \,|\, \mathcal{F}_t^R\right] = a^\top \hat{Y}_t.$$

- In this sense, the HMM with $\sigma_t = a^\top \hat{Y}_t$ is the HMM closest to MSM.

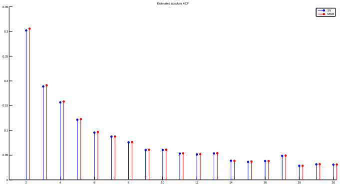
Comparison of some econometric properties

Square distance of HMM with σ_t and MSM with volatility $a^\top Y_t$ is minimized by

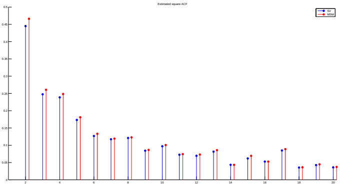
$$\sigma_t = f(\hat{Y}_t) = a^\top \hat{Y}_t.$$



MSM vs. HMM with σ_t



Estimated absolute ACF

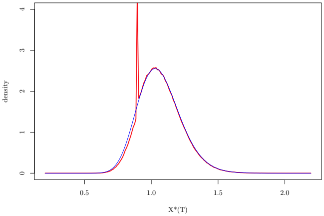


Estimated square ACF

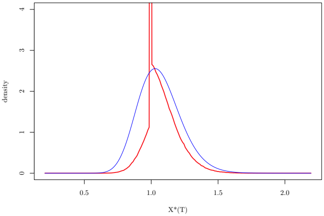
- Conclusion

Model choice, risk constraints and expert opinions

- Model choice: Wrong model might work better in view of estimation errors: In a Black Scholes model with $\mu \in [a, b]$ using an HMM with states a, b outperforms using constant but estimated μ.
- Suitable bounds a, b can be obtained by semi-dynamic risk constraints, see Cuoco/He/Issaenko 2007, Putschögl/S. 2011.
- Static risk constraints on the distribution of the terminal wealth can be included. E.g., for $\varepsilon = 0.01$ and binding constraint $\mathrm{E}[\hat{Z}_T(X_T^* - q)^-] = \varepsilon$:



Pdf of X_T^* without and with risk constraint $q = 0.9$ (atom 2.94%)



Pdf of X_T^* without and with risk constraint $q = 1.0$ (atom 40.21%)

  See Basak/Shapiro 2001, Gabih/S./Wunderlich 2009, S./Wunderlich 2010
- Expert opinions: Frey/Gabih/Wunderlich 2012/14, G./Kondakji/S./W. 2014

Summary and related models

- Differences of HMM and MSM:
  - In HMM: Full and partial information. Partial information with constraints on strategy is consistent approximation for discrete-time model.
  - In MSM: In continuous time only full information. No good approximation for discretized model.
  - But MSM has better econometric properties
  - HMM with non-constant volatility might be a good compromise.
  - Non-constant volatility can be chosen to minimize distance HMM–MSM.
- Filtering, estimation and optimization work for n stocks.
- Similar questions regarding continuous versus discrete-time model for models with Lévy noise with compound Poisson part.
- Other models for μ which allow for explicit filtering and computation of optimal strategies:

  μ as an Ornstein-Uhlenbeck process; leads to Kalman filtering (Lakner 1998).

Further reading

- S./Haussmann (2004) *Optimizing the terminal wealth under partial information: The drift process as a continuous time Markov chain*, Finance and Stochastics 8, 553-577.
- Haussmann/S. (2004) *Optimal terminal wealth under partial information for HMM stock returns*. In: G. Yin and Q. Zhang (eds.): Mathematics of Finance: Proceedings of an AMS-IMS-SIAM Summer Conference June 22-26, 2003, Utah, AMS Contemporary Mathematics 351, 171–185.
- Hahn/Putschögl/S. (2007) *Portfolio optimization with non-constant volatility and partial information*, Brazilian Journal of Probability and Statistics 21, 27–61.
- Elliott/Krishnamurthy/S. (2008) *Moment based regression algorithm for drift and volatility estimation in continuous time Markov switching models*, Econometrics Journal, 11, 244–270.
- Gabih/S./Wunderlich (2009): *Utility maximization under bounded expected loss*, Stochastic Models 25, 375–407.
- S./Wunderlich (2010): *Optimal portfolio policies under bounded expected loss and partial information*, Mathematical Methods of Operations Research 72, 25–61.
- Hahn/Frühwirth-Schnatter/S. (2010) *Markov chain Monte Carlo methods for parameter estimation in multidimensional continuous time Markov switching models*, Journal of Financial Econometrics 8, 88–121.
- Putschögl/S. (2011): *Optimal investment under dynamic risk constraints and partial information*, Quantitative Finance 11, 1547–1564.
- Gabih/Kondakji/S./Wunderlich (2014): Expert opinions and logarithmic utility maximization in a market with Gaussian drift, Communications on Stochastic Analysis 8, 27–47.