Optimizing the Terminal Wealth under Partial Information: The Drift Process as a Continuous Time Markov Chain *

J. Sass, U. Haussmann †

www.math.ubc.ca/~uhaus/prep.html

November 21, 2003

*This work was supported by NSERC under NCE grant 30354 and research grant 88051.

†Department of Mathematics, University of British Columbia, 1984 Mathematics Rd, Vancouver BC, V6T 1Z2, Canada. Email: uhaus@math.ubc.ca, joern.sass@oeaw.ac.at.

MODEL: On $(\Omega, \mathcal{F}, P)$

$$dS_t^0 = S_t^0 \, r_t \, dt, \quad S_0^0 = 1, \; i = 1, \ldots, n,$$

$$dS_t^i = S_t^i(\mu_t^i \, dt + \sum_{j=1}^n \sigma^{ij} \, dW_t^j), \quad S_0^i = s_0^i,$$

$\$\pi_t^i$ invested in stock i at time t,

$\$X_t^\pi$ - total wealth at time t.

Solve

$$\max_\pi E \, U(X_T^\pi)$$

U is a utility function,

r is S adapted, bounded,

σ is non-singular, constant or $\cdots$

μ, W unknown, π must be S adapted

i.e. π must be adapted to the filtration $\{\mathcal{F}_t^S\}$,

N.b. $\mathcal{F}_t^S$ is the information contained in S up to time t.

Lakner: μ is Gaussian, i.e. $d\mu = \alpha(\delta - \mu)dt + \gamma\, d\hat{W}$.

New: state of economy is $Y_t \in \{e^1, e^2, \ldots, e^d\}$ in $\mathbb{R}^d$

Y is a Markov Chain with rate matrix Q:

$$Q_{kl} = \lim_{t \searrow 0} \frac{1}{t} P(Y_t = e_l \,|\, Y_0 = e_k),\ \ k, l = 1 \ldots, d,\ \ k \neq l$$

and $\lambda_k = -Q_{kk} = \sum_{l=1, l\neq k}^{d} Q_{kl}$ the rate of leaving e_k.

$$\mu_t = BY_t,$$

B constant for now.

N.b.

The ith column of B gives μ when the economy is in state i.

If $\quad d = 2 \quad$ i.e. good economy or bad economy, then

$$Q = \begin{pmatrix} -\lambda_1 & \lambda_1 \\ \lambda_2 & -\lambda_2 \end{pmatrix}.$$

Literature

portfolio optimization under partial obs'n

Karatzas and Xue, 1991 - project onto a completely obs. prob.

Lakner, 1995 - basic existence result using filtering, $r = 0$

Lakner, 1998 - π for μ linear Gaussian dynamics, r const.

Zohar, G., 2001 - more explicit calculations for previous model

Pham, H., Quenez, M.-C., 2001 - S-identifiable stoch vol., linear Gaussian dynamics

HMM filters and parameter estimation

Elliott and Rishel, 1994 - estimate μ, n=1

James, M. R., Krishnamurthy, V., Le Gland, F., 1996 - robust filters

Krishnamurthy, V., Elliott, R., 2002 - robust filters

Filtering

Return on ith stock

$$dR_t^i = \frac{dS_t^i}{S_t^i}, \qquad dR_t = \mu_t\,dt + \sigma\,dW_t$$

Excess return:

$$\tilde{R}_t = R_t - \int_0^t r_s \mathbf{1}_n\,ds \quad = \int_0^t \sigma\,d\tilde{W}_s.$$

$$\tilde{W}_t = W_t + \int_0^t \Theta_s\,ds, \text{ a } \tilde{P} \text{ Brownian motion,}$$

$\tilde{P}$ is the equivalent martingale measure with density $\frac{d\tilde{P}}{dP} = Z_T$, and market price of risk: $\Theta_s = \sigma^{-1}\tilde{B}_s Y_s$, $\tilde{B}_t = B - r_t\,\mathbf{1}_{n\times d}$.

HMM filtering (Hidden Markov Model)

$$\eta_t = \mathrm{E}[Y_t \mid \mathcal{F}_t^S], \quad \zeta_t = \mathrm{E}[Z_T \mid \mathcal{F}_t^S], \quad \mathcal{E}_t = \tilde{\mathrm{E}}[Z_T^{-1} Y_t \mid \mathcal{F}_t^S].$$

$$\mathcal{E}_t = \zeta_t^{-1}\,\eta_t, \quad \sum_{i=1}^d \eta_t^i = 1, \quad \sum_{i=1}^d \mathcal{E}_t^i = \zeta_t{}^{-1}$$

$$\boxed{\mathcal{E}_t = \mathrm{E}[Y_0] + \int_0^t Q^\top \mathcal{E}_s\,ds + \int_0^t \mathrm{Diag}(\mathcal{E}_s)\tilde{B}_s^\top(\sigma\sigma^\top)^{-1}\,d\tilde{R}_s}$$

Optimal Trading Strategies

Assumption

$$dr_t = \nu(t, r_t, \mathcal{E}_t)\, dt + \varrho(t, r_t, \mathcal{E}_t)\, d\tilde{R}_t,$$

where ν and ϱ are suitably smooth.

N.b. Interest rate models (e.g. Vasicek, Heath-Jarrow-Morton) are stated w.r.t. $\tilde{P}$.

Set $\beta_t = (S_t^0)^{-1}$, $\tilde{\zeta}_t = \beta_t \zeta_t$,

and $\psi(y)$, $\phi(y)$ are given in terms of $(U')^{-1}$.

Theorem Under some technical assumptions the optimal trading strategy is

$$\begin{aligned} \hat{\pi}_t \;=\; & \frac{\beta_t^{-1}}{\hat{y}} (\sigma\, \sigma^\top)^{-1} \Big\{ \tilde{B}\, \mathcal{E}_t\, \tilde{\mathrm{E}}[\psi(\hat{y}\, \tilde{\zeta}_T) \mid \mathcal{F}_t^S] \\ & + \tilde{\mathrm{E}}[\psi\, (\hat{y}\, \tilde{\zeta}_T) \int_t^T ((\sigma\, D_t \mathcal{E}_s) \tilde{B}^\top - (\sigma\, D_t r_s) \zeta_s^{-1} \mathbf{1}_n^\top)(\sigma\, \sigma^\top)^{-1} d\tilde{R}_s \mid \mathcal{F}_t^S] \\ & + \tilde{\mathrm{E}}[\hat{y}\, \varphi(\hat{y}\, \tilde{\zeta}_T) \beta_T \int_t^T (\sigma\, D_t r_s)\, ds \mid \mathcal{F}_t^S] \Big\}. \end{aligned}$$

$\hat{y}$ is Lagrange multiplier (budget equation),
D_t is Malliavin derivative.

mean-variance hedge + hedge for fluctuations in market price of risk, Θ, + hedge for other fluctuations in interest rates.

Log and Power Utility

Corollary

For logarithmic utility $U(x) = log(x)$,

$$\boxed{\hat{\pi}_t = (\sigma\sigma^\top)^{-1}\tilde{B}_t\eta_t\hat{X}_t.}$$

and for power utility $U(x) = \frac{x^\alpha}{\alpha}$, $\alpha < 1$, $\alpha \neq 0$,

$$\begin{aligned}\hat{\pi}_t = \\ &\frac{\hat{X}_t(\sigma\sigma^\top)^{-1}}{(1-\alpha)\tilde{\mathrm{E}}[\beta_{t,T}\tilde{\zeta}_{t,T}^{\frac{1}{\alpha-1}} \mid r_t, \mathcal{E}_t]}\{\tilde{B}_t\eta_t\tilde{\mathrm{E}}[\tilde{\zeta}_{t,T}^{\frac{\alpha}{\alpha-1}} \mid r_t, \mathcal{E}_t] \\ &+ \tilde{\mathrm{E}}[\tilde{\zeta}_{t,T}^{\frac{\alpha}{\alpha-1}} \int_t^T ((\sigma D_t\mathcal{E}_{t,s})\tilde{B}_s^\top - (\sigma D_t r_s)\zeta_{t,s}^{-1}\mathbf{1}_n^\top)(\sigma\sigma^\top)^{-1}\, d\tilde{R}_s \mid r_t, \mathcal{E}_t] \\ &+ \alpha\, \tilde{\mathrm{E}}[\beta_{t,T} \int_t^T (\sigma D_t r_s)\, ds \mid r_t, \mathcal{E}_t]\},\end{aligned}$$

where $\beta_{t,T} = \beta_T/\beta_t$, $\tilde{\zeta}_{t,T} = \tilde{\zeta}_T/\tilde{\zeta}_t$, $\zeta_{t,s}^{-1} = \zeta_s^{-1}/\zeta_t^{-1}$ and $\mathcal{E}_{t,s} = \mathcal{E}_s/\zeta_t^{-1}$.

Parameter Estimation - $\sigma\sigma^\top$

$\sigma\sigma^\top = [R]_t/t = \lim_{h\to 0}\sum(\Delta_h R)(\Delta_h R)^\top$, but $h = 1$ day. Instead

$$\begin{aligned}\frac{1}{h}\,\mathrm{E}[R_h R_h^\top] &= \sigma\sigma^\top + \frac{1}{h}\,E[\int_0^h \mu_t\,dt \int_0^h \mu_s^\top\,ds] \\ &\approx \sigma\sigma^\top + c_1 h + c_2 h^2 + \dots\end{aligned}$$

Y stationary so left side is

$$\frac{1}{h}\mathrm{E}[R_h R_h^\top] \approx \frac{1}{T}\sum_{k=1}^{T/h}(R_{kh} - R_{(k-1)h})(R_{kh} - R_{(k-1)h})^\top$$

Algorithm for $\sigma\sigma^\top$

(1) Estimate $\frac{1}{h}\,E[R_h R_h^\top]$, $h = \Delta t, 2\,\Delta t \dots, m\,\Delta t$.

(2) Find the least square quadratic fit to

$$(h, \frac{1}{h}\,E[R_h R_h^\top])_{h=\Delta t,\dots,m\,\Delta t}$$

(3) Choose the y-intercepts of the regression curves as estimate for $\sigma\sigma^\top$.

Parameter Estimation - B,Q

Knowing $\sigma\sigma^\top$ we use the EM algorithm to estimate B and Q.

A sequential procedure for maximizing a Likelihood function.

Requires unnormalized filters of the form

$$\mathcal{E}_t(X) = \tilde{\mathrm{E}}[Z_t^{-1} X_t \mid \mathcal{F}_t^S]$$

for various X, eg occupation times $O^k = \int_0^T Y_s^k \, ds$. These are found as solutions to simple linear SDE's driven by $\tilde{R}$ and then reformulated as robust filters:

Φ_t is a stochastic fundamental matrix, depends on $\tilde{B}$,

$\bar{\mathcal{E}}_t = \Phi_t^{-1} \mathcal{E}_t$,

Need

$$\begin{aligned}
d\bar{\mathcal{E}}_t &= \Phi_t^{-1} Q^\top \Phi_t \bar{\mathcal{E}}_t dt, \qquad \bar{\mathcal{E}}_0 = E(Y_0), \\
d\bar{\mathcal{E}}_t(O^k Y) &= (\Phi_t^{-1} Q^\top \Phi_t \bar{\mathcal{E}}_t(O^k Y) + \bar{\mathcal{E}}_t^k e_k) dt, \\
d\bar{\mathcal{E}}_t(\tilde{O}^k Y) &= (\Phi_t^{-1} Q^\top \Phi_t \bar{\mathcal{E}}_t(\tilde{O}^k Y) + r_t \bar{\mathcal{E}}_t^k e_k) dt, \\
d\bar{\mathcal{E}}_t(N^{kl} Y) &= (\Phi_t^{-1} Q^\top \Phi_t \bar{\mathcal{E}}_t(N^{kl} Y) + (\Phi_t^{-1} Q^\top \Phi_t)_{lk} \bar{\mathcal{E}}_t^k e_l) dt, \\
d\bar{\mathcal{E}}_t(G^{k,i} Y) &= \Phi_t^{-1} Q^\top \Phi_t \bar{\mathcal{E}}_t(G^{k,i} Y) \, dt + \bar{\mathcal{E}}_t^k e_k ((\sigma\sigma^\top)^{-1} d\tilde{R}_t)_i.
\end{aligned}$$

Use Euler (=Milstein) approximation to solve.

Figure 1: $\hat{\pi}/\hat{X}$ for logarithmic and power utlity ($\alpha = 0.5, -5$)

Simulation - no parameter identification

$r = 0.06$, $n = 1$ stock, 2 states $B = (b_1, b_2)$, $\quad b_1 > b_2$.

	σ	b_1	b_2	λ_1	λ_2
parameters	0.20	0.80	-0.40	30	24

$$U(x) = \log(x), \qquad U(x) = x^\alpha/\alpha,\ \alpha = 0.5,\ -5.$$

50 simulations. Note the extreme long and short positions!

Simulation - one stock, parameter identification

Simulate 6 years (250 trading days each).

Use first 5 to estimate parameters.

Then optimize portfolio over last year.

Use $m = 4, \Delta t = 1/250$ (i.e. 1 day) for $\sigma\sigma^\top$.

Average the $5 \times 250/4$ estimates of $\sigma\sigma^\top$.

Initialize EM with $b_1 = \bar{\mu} + 0.5,\ b_2 = \bar{\mu} - 0.5$ where

$\bar{\mu}$ is the average return per unit time over the 5 years.

$\lambda_1 = 34,\ \lambda_2 = 28$. Iterate EM 5 times!

	σ	b_1	b_2	λ_1	λ_2
true parameters	0.20	0.80	-0.40	30	24
estimated par.	0.200	0.771	-0.412	33.05	28.36
standard dev.	0.008	0.22	0.20	1.33	1.12
abs. error in %	0.01	3.58	3.04	10.2	18.2

Table 1: Estimation of parameters for 1 stock with 2 states

Simulation - one stock, optimization

Take initial wealth $x_0 = 1$.

	$U(x)$	$\log(x)$			$-x^{-5}/5$		
	strategy	opt	Mert	b/h	opt	Mert	b/h
est.	av. $\hat{X}_T$	4.60	1.37	1.16	1.32	1.09	1.16
par.	av. $U(\hat{X}_T)$	0.104	0.001	0.118	< -100	-0.165	-0.214
	med. $U(\hat{X}_T)$	0.130	0.060	0.126	-0.094	-0.142	-0.107
	opt better than		291	256		347	286
known	av. $\hat{X}_T$	3.36	1.25	1.16	1.22	1.09	1.16
par.	av. $U(\hat{X}_T)$	0.399	0.136	0.118	-0.121	-0.141	-0.214
	med. $U(\hat{X}_T)$	0.305	0.150	0.126	-0.091	-0.131	-0.107
	opt better than		296	288		359	292

Table 2: Wealth and utility for 1 stock with 2 states

Mert means "Merton" strategy: $\pi_t = \left(\frac{1}{1-\alpha}\right) \frac{\bar{\mu}-r}{\sigma^2} \hat{X}_t$

$$\frac{1}{N} \sum_{i=1}^{N} \Delta R_{i\Delta t} \approx E R_{\Delta t} = \bar{\mu} \Delta t, \quad \Delta t = 1.$$

b/h: buy the stock only and hold to time T

Historical prices

20 stocks, 30 years (1972 - 2001),

each stock gives rise to 25 experiments lasting 6 years:

years 1-6, 2-7, ... 25-30 (inclusive). (Not independent!)

So 500 experiments for "one stock" each lasting 6 years -

identify parameters for 5 and optimize over last year.

Interest rate is fed rate.

For optimization, take $D_t r_s = 0$

(investor uses constant interest rate, avg. of the previous year,

but he is exposed to market rate, i.e. in $\hat{X}$).

Only one iteration of EM (slow PC)

The average estimated parameters were

$$\sigma \approx 0.26, \quad b_1 \approx 0.66, \quad b_2 \approx -0.37, \quad \lambda_1 = \lambda_2 \approx 125.8$$

outliers

$U(x)$		$\log(x)$			$-x^{-5}/5$	
strategy	opt	Mert	b/h	opt	Mert	b/h
av. $\hat{X}_T$	1.509	1.134	1.149	1.129	1.096	1.149
med. $\hat{X}_T$	1.160	1.090	1.118	1.107	1.088	1.118
av. $U(\hat{X}_T)$	0.012	-0.030	0.110	-538.3	-0.143	-0.454
med. $U(\hat{X}_T)$	0.125	0.083	0.111	-0.121	-0.131	-0.115
aborted	11	3		0	0	
opt better than		288	258		303	243

Table 3: Wealth and utility for historical prices, 1 stock with 2 states

More historical

log utility

setting		I		II		III	
strategy	b/h	opt	Mert	opt	Mert	opt	Mert
av. $\hat{X}_T$	1.149	1.509	1.134	1.594	1.160	1.448	1.160
med. $\hat{X}_T$	1.118	1.160	1.090	1.154	1.093	1.134	1.093
av. $U(\hat{X}_T)$	0.124	0.012	-0.030	0.037	-0.007	0.034	-0.007
med. $U(\hat{X}_T)$	0.111	0.125	0.083	0.124	0.082	0.100	0.082
aborted		11	3	12	2	12	2
opt better than			288		289		281

Table 4: I: 2 states, const. r; II: 2 states, historical r; III: 3 states, historical r

Extensions (Stochastic volatility)

We can extend theoretical results to σ and B being S adapted processes, i.e.

$r,\ \sigma,\ \sigma^{-1}\tilde{B}$ are **known** smooth functions of (t, S_t, η_t).

Applied to historical prices with B constant and $\sigma_t = s_0 + s_1\eta_t^1 + s_2(\eta_t^1)^2$, s_i to be estimated, get better results:

	opt	const σ	Merton	b/h
av. $\hat{X}_T$	1.745	1.603	1.163	1.153
med. $\hat{X}_T$	1.181	1.138	1.095	1.121
av. $\log(\hat{X}_T)$	0.122	0.057	0.006	0.116
med. $\log(\hat{X}_T)$	0.167	0.129	0.091	0.114
aborted	1	11	2	-

Table 5: 2 states, historical r, stochastic vol.

"const σ" is the strategy where σ is assumed to be constant, and is estimated as before.

Merton strategy also assumes constant σ and μ.