Some Applications of M-ary Detection in Quantitative Finance

W. P. Malcolm[(1,2)] & R. J. Elliott[3]

(1) **Statistical Machine Learning Program, National ICT, Australia**

Email: `paul.malcolm@nicta.com.au/paul.malcolm@anu.edu.au`

(2) **Mathematical Sciences Institute, ANU**

(3) **Haskayne School of Business, University of Calgary, Canada**

Email: `relliott@ucalgary.ca`

Literature Background:

1. *"Reproducing Gaussian Densities"*, Elliott, Malcolm, Systems and Control Letters 2000.
2. *"M-ary Detection for Gauss-Markov Models"*, Elliott, Malcolm, IEEE Trans. AC 2005.
3. *"Noise Gain Detection"* Malcolm, Bensoussan, submitted SIAM J. Opt. & Control. 2008.
4. *"Mean-Annual-Return Estimation"* Malcolm, Elliott, (accepted), J. Quantitative Finance.
5. *"Pairs Trading"* van der Hoek, Elliott and Malcolm, J. Quantitiative Finance, 2004
6. *"Extensions to Filter-based Pairs Trading"* Malcolm, Elliott, (submitted to) J. Quantitative Finance

Seminar Plan:

1. *What is M-ary Detection in our context ?*
2. *M-ary Detection applied to mean-return estimation*
3. *M-ary Detection applied to volatility estimation, a HMM example*
4. *M-ary Detection applied to a pairs trading scheme*
5. *Future work*

Some Background

- The term M-ary detection has its origin in Electrical Engineering, however, in statistics, essentially the same technique is referred to as sequential analysis (Abraham Wald), or more simply, sequential hypothesis testing.
- Our particular interest will be to test for "a priori" specified candidate model-parameter-sets for a given model form.
- Model calibration is central to many areas of estimation, the most common approach being parameter estimation via the so called EM algorithm.
- Our motivation is to provide a practical alternative to the EM algorithm which can be slow to converge and suffer numerical issues.

DEFINITION 1 *In our context, solving an M-ary detection problem will mean estimating probabilities of the form $\widehat{P}_{\mathbf{time}}(Model \mid Observation_{\mathbf{time}})$.*

Let us recall typical assumptions in basic state estimation:

Consider the following ubiquitous stochastic dynamical system, defined on some space $(\Omega, \mathcal{F}, P)$:

$$x_{k+1} = A_k x_k + B_k w_k, \quad \text{(State Model)} \tag{1}$$

$$y_k = C_k x_k + D_k v_k. \quad \text{(Measurement/Observer Model)} \tag{2}$$

Write,

$$\widehat{x}_{k|k} \triangleq E^P \Big[x_k \mid \sigma\{y_\ell, 0 \geq \ell \leq k\} \Big] \tag{3}$$

In computing the state estimate at (4), it is implicitly assumed that the model parameter set, viz, $\{A, B, C, D\}$ is correct. This assumption is almost surely an **incorrect assumption**.

M-ary detection, in a sense, reformulates the parameter estimation over a finite set of M candidate sets of Model hypotheses, $H_j = \{A_j, B_j, C_j, D_j\}$, where $j \in \{1, 2, \ldots, M\}$.

One wishes to estimate the probabilities $\widehat{p}_k^j \triangleq P(H = H_j \mid \sigma\{y_\ell, 0 \geq \ell \leq k\})$.

In what follows we will recast our estimation problem to include and uncertain model, that is, formally, we consider :

$$\widehat{x}_{k|k} \stackrel{\Delta}{=} E^P\Big[x_k \mid \sigma\{y_\ell, 0 \geq \ell \leq k\} \,\&\, \text{Model}\Big] \tag{4}$$

Now the model is explicitly included

Mean-Return Detection

We shall consider variations on the following standard log-normal asset price dynamics,

$$S_t = S_0 + \int_0^t \mu S_u du + \int_0^t \sigma S_u dB_u. \tag{5}$$

Here, S is the unit price of an asset, μ is the asset price mean-annual-return, σ is the asset volatility parameter and B is a standard Brownian motion.

We shall use an indicator function representation for an M-valued simple random variable $\boldsymbol{\alpha}$. The SRV $\boldsymbol{\alpha}$ takes values in $\mathcal{S} = \{\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_M\}$, whose elements $\boldsymbol{f}_j$ are column vectors with unity in the j^{th} position and zero elsewhere,

$$\boldsymbol{f}_j \in \left\{ \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \right\} = \mathcal{L} \subseteq \mathbb{R}^M. \tag{6}$$

We suppose that the true, but unknown mean-return $\boldsymbol{\mu}$, takes one of M possible states. Write

$$\boldsymbol{\mu} \triangleq \big(\langle \boldsymbol{\mu}, \boldsymbol{f}_1 \rangle, \ldots, \langle \boldsymbol{\mu}, \boldsymbol{f}_M \rangle\ \big)' \in \mathbb{R}^M_+. \tag{7}$$

In this scenario we assume the volatility is known in advance, consequently the asset price dynamics are

$$S_t = S_0 + \int_0^t \langle \boldsymbol{\alpha}, \boldsymbol{\mu} \rangle S_u du + \int_0^t \sigma S_u dB_u. \tag{8}$$

While S is observered directly at each sampling epoch it is conveninet to work with the related process $\Gamma_t \triangleq ln\big(S_t\big)$. This new process has dynamics

$$\Gamma_t = \Gamma_0 + \int_0^t \Big(\langle \boldsymbol{\alpha}, \boldsymbol{\mu} \rangle - \tfrac{1}{2}\sigma^2\Big) du + \int_0^t \sigma dB_u. \tag{9}$$

Further, writing $y_t \stackrel{\Delta}{=} \Gamma_t/\sigma = \ln(S_t)/\sigma$, we obtain

$$y_t = y_0 + \int_0^t \langle \boldsymbol{\alpha}, \boldsymbol{\xi} \rangle du + B_t. \tag{10}$$

Here

$$\boldsymbol{\xi} \stackrel{\Delta}{=} \Big(\sigma^{-1}\big(\langle \boldsymbol{\mu}, \boldsymbol{f}_1 \rangle - \tfrac{1}{2}\sigma^2\big), \sigma^{-1}\big(\langle \boldsymbol{\mu}, \boldsymbol{f}_2 \rangle - \tfrac{1}{2}\sigma^2\big), \ldots, \sigma^{-1}\big(\langle \boldsymbol{\mu}, \boldsymbol{f}_M \rangle - \tfrac{1}{2}\sigma^2\big)\Big)'. \tag{11}$$

The M-ary detection problem we wish to solve, recursively, is to compute the expectations,

$$E\Big[\mathbf{1}_{\{\omega:\, \boldsymbol{\alpha}(\omega) = \boldsymbol{f}_i\}} \mid \sigma\{y_u, 0 \le u \le t\}\Big], \quad \text{for} \quad i = 1, 2, \ldots, M. \tag{12}$$

In fact the problem just stated can be solved exactly when considered as a <u>degenerate Wonham filter</u>.

Notation: We define our filtrations (information records) as follows:

$$\mathbb{Y} = \big\{\mathcal{Y}_{0,t}\big\}, \text{ with } \quad \mathcal{Y}_{0,t} = \sigma\big\{y_u, u \le t\big\} \tag{13}$$

$$\mathbb{G} = \big\{\mathcal{G}_{0,t}\big\}, \text{ with } \quad \mathcal{G}_{0,t} = \sigma\big\{\boldsymbol{\alpha}, y_u, u \le t\big\} \tag{14}$$

To compute each of the M detector probabilites corresponding to the individual model hypotheses, we consider the expectations $E\big[\big\langle \boldsymbol{\alpha}, \boldsymbol{f}_j \big\rangle \mid \mathcal{Y}_{0,t}\big]$. In what follows, we shall choose to do our calculations under a new, ("reference"), probability measure $P^\dagger$. Under the measure P, we have dynamics of the form

$$P: \quad \begin{cases} \boldsymbol{\alpha} = \boldsymbol{\alpha}, \\ y_t = y_0 + \displaystyle\int_0^t \big\langle \boldsymbol{\alpha}, \boldsymbol{\xi} \big\rangle \, du + B_t. \end{cases} \tag{15}$$

Under $P^\dagger$, our dynamics are:

$$P^\dagger: \quad \begin{cases} \boldsymbol{\alpha} = \boldsymbol{\alpha}, \\ y = \text{``is a standard Brownian motion''} \end{cases} \tag{16}$$

The Radon-Nikodym derivative, from $P^\dagger$, to P, is denoted by Λ_t, where

$$\Lambda_t \triangleq \frac{dP}{dP^\dagger}\Big|_{\mathcal{G}_{0,t}} = \exp\Big\{ \big\langle \boldsymbol{\alpha}, \boldsymbol{\xi} \big\rangle y_t - \tfrac{1}{2}\big\langle \boldsymbol{\alpha}, \boldsymbol{\xi} \big\rangle^2 t \Big\}. \tag{17}$$

Then Λ is $(P, \mathcal{G})$-martingale and has dynamics

$$\Lambda_t = 1 + \big\langle \boldsymbol{\alpha}, \boldsymbol{\xi} \big\rangle \int_0^t \Lambda_u dy_u. \tag{18}$$

Write

$$q_t^j \triangleq E^\dagger\big[\Lambda_t \big\langle \boldsymbol{\alpha}, \boldsymbol{f}_j \big\rangle \mid \mathcal{Y}_{0,t}\big]. \tag{19}$$

The scalar-valued quantity q_t^j, is the estimated un-normalised probability that model H_j explains the observation process.

Theorem 1 *The un-normalised scalar-valued probability q^j, defined at (19), has dynamics*

$$q_t^j = q_0^j + \langle \boldsymbol{f}_j, \boldsymbol{\xi} \rangle \int_0^t q_u^j dy_u. \tag{20}$$

The quantities q^j are each determined individually and each depend upon the observed data process y. However, these estimated probabilities are combined through normalisation. The final estimated, (normalised), probabilities are computed as

$$\widehat{p}^j = P(H = H_j \mid \mathcal{Y}_{0,t}) = \frac{q_t^j}{\sum_{\ell=1}^M q_t^\ell} \tag{21}$$

The dynamics at (20) include a stochastic intergral. However, these dynamics have an exact solution due to a Theorem of Doleans-Dade, which we now recall.

THEOREM 2 (DOLEANS-DADE) *Suppose Z is a semimartingale and suppose $Z_{0-} = 0$ $P - a.s.$ Then there exists a unique semimartingale q, such that*

$$q_t = q_{0-} + \int_{[0,t]} q_{u-} dZ_u. \tag{22}$$

Here

$$q_t = q_{0-} \exp\left\{Z_t - \tfrac{1}{2}\langle Z^c, Z^c\rangle_t\right\} \prod_{0 \le u \le t} (1 + \Delta Z_u) \exp\{-\Delta Z_u\} \tag{23}$$

The exact solution for our detector is given by

$$q_t^j = q_0^j \exp\left\{\big\langle \boldsymbol{\xi}, \boldsymbol{f}_j \big\rangle y_t - \tfrac{1}{2}\big\langle \boldsymbol{\xi}, \boldsymbol{f}_j \big\rangle^2 t\right\}. \tag{24}$$

Here,

$$\big\langle \boldsymbol{\xi}, \boldsymbol{f}_j \big\rangle = \Big(\big\langle \boldsymbol{\mu}, \boldsymbol{f}_j \big\rangle - \tfrac{1}{2}\sigma^2\Big)/\sigma. \tag{25}$$

Simulation Example

To investigate the performance of mean-return Mary detection, we consider a single realisation of the asset price model at (5), with true parameter values $\mu = 0.1$ and $\sigma = 0.2$. This realisation is shown in Figure 1. Since the mean-return is unbounded, (in theory), and can take both positive and negative values, we take as our set of candidate model hypotheses is :

$$\boldsymbol{\mu} \stackrel{\Delta}{=} (-0.6 - 0.5, -0.4, -0.3, -0.2, -0.1,\ 0.1,\ 0.2,\ 0.3,\ 0.4,\ 0.5,\ 0.6)'. \qquad (26)$$

Remark 1 Detection allows one to bound the range of a parameter, this is generally not an easy task with the EM algorithm.

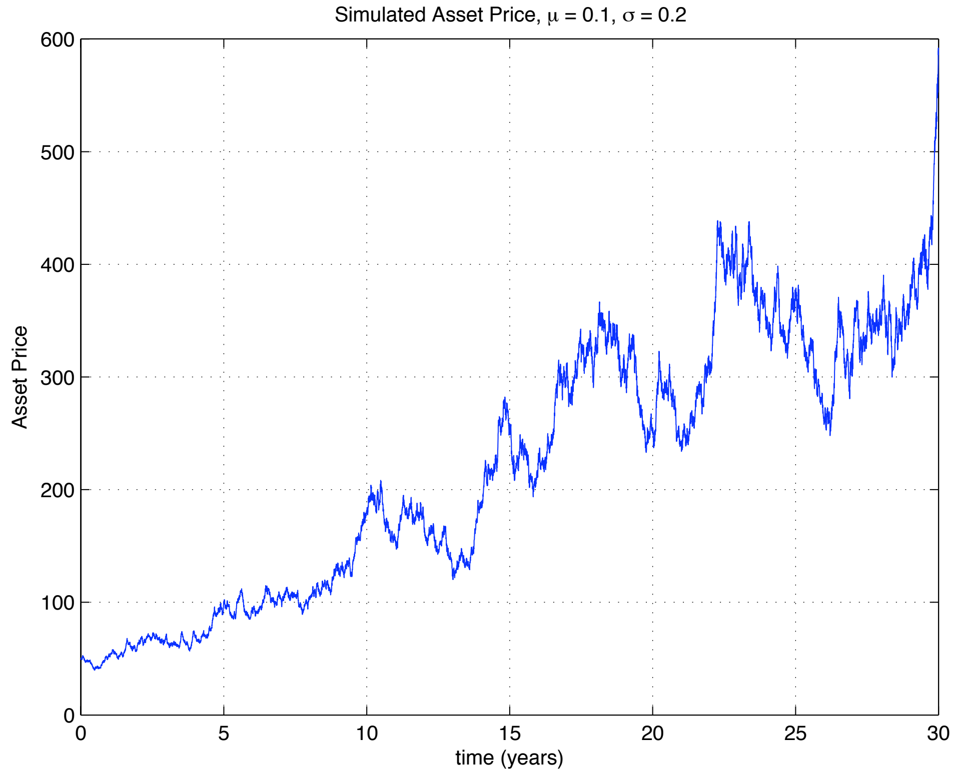

Figure 1: A single realisation of the asset price model described by equation (5), mean-return $\mu = 0.1$, volatility $\sigma = 0.2$. The dependent variable is in an arbitrary currency.

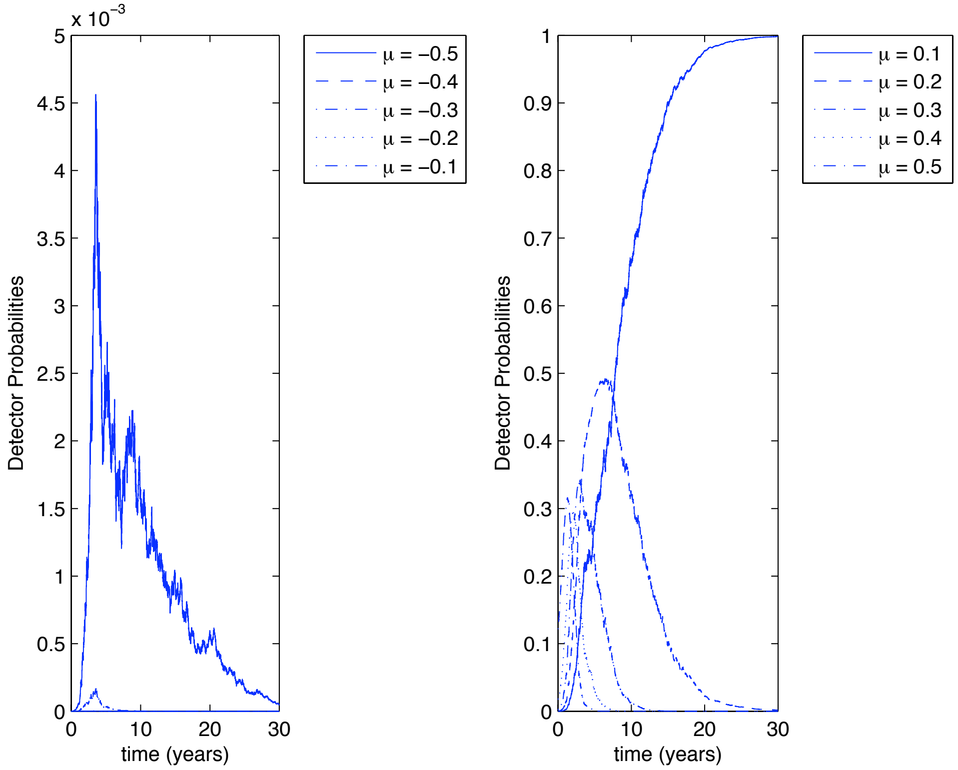

Figure 2: Here we show the evolution of four estimated M-ary detector probabilities corresponding to a collection of candidate mean-returns. The dominant curve which converges to unity corresponds to the mean return value $\mu = 0.1$.

REMARK 2 It is well known that the estimation of mean-return is an exceptionally difficult problem. Some key papers considering this problem are:

- Black, F., *Estimating Expected Return*, Financial analysis Journal, January-February 1995, pp. 168-171
- Merton, Robert J. *On estimating the expected return on the market*, Journal of Financial Economics, 8 (1980), pp. 323-361.
- Goldenberg, D. H. and Schmidt, J. R., *On estimating the expected rate of return in diffusion price models with application to estimating the expected return on the market*, The journal of Financial Quantitive Analysis, Volume 31, Numner 4, December 1996, pp. 605-631

The article above by **Goldenberg and Schmidt** is potentially of most interest to the practitioner. In this article the conventional Maximum Likelihood (ML) estimator is tested to estimate the mean return for the same model in the simulation above. It is shown that the ML estimator takes more than 120 years of data to converge.

HMM Example:

Suppose that the volatility $\boldsymbol{\sigma}$ takes values in the set $\{\sigma_1, \ldots, \sigma_n\}$ and is determined by the state of a Markov chain X evolving in discrete time. We take the state space of our Markov chain to be a canonical basis of orthonormal unit vectors, that is, a basis with elements

$$\boldsymbol{e}_\ell \triangleq (0, 0, \ldots, 1, \ldots, 0)', \tag{27}$$

where,

$$\boldsymbol{e}_\ell \in \left\{ \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \right\} = \mathcal{S} \subseteq \mathbb{R}^n. \tag{28}$$

Further

$$\pi_{(j,i)} \triangleq P(X_{k+1} = \boldsymbol{e}_j \mid X_k = \boldsymbol{e}_i) \tag{29}$$

$$= P(X_1 = \boldsymbol{e}_j | X_0 = \boldsymbol{e}_i). \tag{30}$$

To denote the matrix of transition probabilities for the process X, we write

$$\Pi = \left[\pi_{(j,i)}\right]_{\substack{1\leq j\leq n \\ 1\leq i\leq n}}. \tag{31}$$

Write $\mathcal{F}_k \stackrel{\Delta}{=} \sigma\{X_0, X_1, \ldots, X_k\}$.
Then

$$X_{k+1} = \Pi X_k + L_{k+1} \quad \text{(here } L \text{ is a } (\mathcal{F}, P)\text{-martingale increment).} \tag{32}$$

We shall suppose that the above Markov chain appears as a discretisation of the continous-time model (5) and acts as a modulating process, switching the volatility value, according to the state of X.

$$S_{k+1} = S_k \exp\Big\{\mu\Delta_t - \tfrac{1}{2}\int_k^{k+\Delta_t} \sigma_u^2 du + \int_k^{k+\Delta_t} \sigma_u dB_u\Big\}. \tag{33}$$

Write

$$\begin{aligned} Y_{k+1} \stackrel{\Delta}{=} \log\Big\{\frac{S_{k+1}}{S_k}\Big\} &= \mu\Delta_t - \tfrac{1}{2}\int_k^{k+\Delta_t} \sigma_u^2 du + \int_k^{k+1} \sigma_u dB_u \\ &\approx \mu\Delta_t - \tfrac{1}{2}\big\langle \boldsymbol{\sigma}, X_k\big\rangle^2 \Delta_t + \big\langle \boldsymbol{\sigma}, X_k\big\rangle \sqrt{\Delta_t} w_{k+1}. \end{aligned} \tag{34}$$

Here $\{w_1, w_2, \ldots, w_k, \ldots\}$ is a sequence of independent and identically Gaussian $N(0, 1)$ random variables. Write $\mathcal{Y}_k = \sigma\{Y_1, Y_2, \ldots, Y_k\}$.

Remark 3 $\{Y\}$ is a sequence of logarithmic price increments and is computed exactly. Our approximation is in the dynamics explaining the evolution of Y.

The new dynamical system we now consider is

$$X_{k+1} = \Pi X_k + L_{k+1} \tag{35}$$

$$Y_{k+1} = a - \langle \boldsymbol{b}, X_k \rangle + \langle \boldsymbol{c}, X_k \rangle w_{k+1}. \tag{36}$$

Here,

$$a \triangleq \mu \Delta_t \tag{37}$$

$$\boldsymbol{b} \triangleq \frac{\Delta_t}{2} \left(\langle \boldsymbol{\sigma}, \boldsymbol{e}_1 \rangle^2, \ldots, \langle \boldsymbol{\sigma}, \boldsymbol{e}_n \rangle^2 \right)' \tag{38}$$

$$\boldsymbol{c} \triangleq \sqrt{\Delta_t} \left(\langle \boldsymbol{\sigma}, \boldsymbol{e}_1 \rangle, \ldots, \langle \boldsymbol{\sigma}, \boldsymbol{e}_n \rangle \right)'. \tag{39}$$

To denote a single model hypothesis for our discrete-time dynamics we write

$$H_j \triangleq \{\Pi^{H_j}, a^{H_j}, \boldsymbol{b}^{H_j}, \boldsymbol{c}^{H_j}\}, \quad j \in \{1, 2, \ldots, M\}. \tag{40}$$

Using a simple random variable $\boldsymbol{\alpha}$, we now wish to evaluate the expectation

$$q_k^j = E^\dagger\big[\Lambda_k \langle \boldsymbol{\alpha}, \boldsymbol{f}_j \rangle \mid \mathcal{Y}_k\big]. \tag{41}$$

Here $\boldsymbol{\alpha} = \boldsymbol{f}_j$ iff $H = H_j$.

To make a clear distinction between the filter information-state defined for specific model H_j, and the corresponding unnormalised detector probability for model H_j, we write, respectively

$$q_k^{H_j} \triangleq E^\dagger\big[\Lambda_k \langle \boldsymbol{\alpha}, \boldsymbol{f}_j \rangle X_k \mid \mathcal{Y}_k\big] \in \mathbb{R}^M, \tag{42}$$

$$q_k^{\mathrm{Det}} \triangleq \Big(E^\dagger\big[\Lambda_k \langle \boldsymbol{\alpha}, \boldsymbol{f}_1 \rangle \mid \mathcal{Y}_k\big], \ldots, E^\dagger\big[\Lambda_k \langle \boldsymbol{\alpha}, \boldsymbol{f}_M \rangle \mid \mathcal{Y}_k\big]\Big)' \in \mathbb{R}^M. \tag{43}$$

THEOREM 3 (M-ARY DETECTION FILTER) *The M-ary detection filter for the model hypothesis H_j is computed by the recursion*

$$\left\langle q_{k+1}^{Det}, \boldsymbol{f}_j \right\rangle = \sum_{\ell=1}^{M} \frac{\Phi\Big(\big(Y_{k+1} - a^{H_j} - \left\langle \boldsymbol{b}^{H_j}, \boldsymbol{e}_\ell \right\rangle\big) / \left\langle \boldsymbol{c}^{H_j}, \boldsymbol{e}_\ell \right\rangle\Big)}{\left\langle \boldsymbol{c}^{H_j}, \boldsymbol{e}_\ell \right\rangle \Phi(Y_{k+1})} \left\{ \frac{\left\langle q_{k-1}^{H_j}, \boldsymbol{e}_\ell \right\rangle}{\left\langle q_{k-1}^{H_j}, \boldsymbol{1} \right\rangle} \right\} \left\langle q_k^{Det}, \boldsymbol{f}_j \right\rangle. \tag{44}$$

REMARK 4 We omit the proof of the above Theorem, however, a key and simple identity used our proof is the following:

$$\begin{aligned} P\big(\text{State} \cap \text{Model} \mid \text{Observations}\big) &= \frac{P\big(\text{State} \cap \text{Model} \cap \text{Observations}\big)}{P\big(\text{Observations}\big)} \\ &= P\big(\text{State} \mid \text{Model} \cap \text{Observations}\big) \times \\ &\qquad P\big(\text{Model} \mid \text{Observations}\big). \end{aligned}$$

Pairs trading Example

Filter-based pairs trading was introduced in the article

Pairs Trading, R. J. Elliott, J. van der Hoek and W. P. Malcolm
Quantitative Finance, Volume 5, Number 3, June 2005

The approach taken in the articke, is to assume the price difference of two similar stocks is described, (in continuous-time), by a mean-reverting Stochastic Differential Equation (SDE). This SDE is given below.

$$dx_t = (a - bx_t)dt + \sigma dB_t \tag{45}$$

Here B is a standard Brownian motion and $a \in \mathbb{R}$, $b > 0$ and $\sigma \geq 0$. The dynamics at (45) may be discretised in a variety of ways. Suppose we consider a first-order Euler discretation at regular sampling points, each separated in time by Δ_t, that is,

$$x_{t_k} - x_{t_{k-1}} = (a - bx_{t_{k-1}})\Delta_t + \sigma\sqrt{\Delta_t}w_{t_k}. \tag{46}$$

Here w is a standard iid $N(0,1)$ Gaussian process.

Write

$$e \triangleq a\Delta_t, \tag{47}$$

$$a \triangleq 1 - b\Delta_t, \tag{48}$$

$$b \triangleq \sigma\sqrt{\Delta_t}. \tag{49}$$

$$\tag{50}$$

Using these identities and writing $x_k = x_{t_k}$, we may write the discrete-time dynamics at (51) as

$$x_k = e + ax_{k-1} + cw_k. \tag{51}$$

We now assume that the process x is not observed directly, rather it is observed through a noisy linear map, that is

$$y_k = x_{k-1} + Dv_k. \tag{52}$$

Here the process v is an i.i.d. $N(0,1)$ Gaussian process.

The idea here is that the additive noise in equation (52) serves to represent/model any penalty for the approximation at (51).

Our discrete-time model is divided into a hidden state process that represents the spread and an observation process that represents a noisy observation of the spread. The hidden state dynamics are :

$$x_{k+1} = e + ax_k + bw_{k+1} \tag{53}$$

$$y_k = x_k + dv_k \tag{54}$$

The estimation problem we wish to solve is to estimate the one-step-ahead prediction of process x, given the process y. We start by writing the dynamics at (53) and (54) as

$$x_{k+1} = \sum_{j=1}^{M} \langle \boldsymbol{\alpha}, \boldsymbol{f}_j \rangle e_j + \sum_{j=1}^{M} \langle \boldsymbol{\alpha}, \boldsymbol{f}_j \rangle a_j x_k \sum_{j=1}^{M} \langle \boldsymbol{\alpha}, \boldsymbol{f}_j \rangle b_j v_{k+1} \tag{55}$$

$$y_k = x_k + \sum_{j=1}^{M} \langle \boldsymbol{\alpha}, \boldsymbol{f}_j \rangle d_j w_k. \tag{56}$$

Here our model parameter-set hypotheses of interest are of the form :
$H_j \triangleq \{e_j, a_j, b_j, d_j\}$ with $e_j, a_j, b_j, d_j \in \mathbb{R}$.

The M-ary detector for Gauss-Markov dynamics is given in the next Theorem.

Remark 5 To simplify the implementation task of coding M separate Kalman filters, as is required by our M-ary detection scheme, one can equivalently code a single higher dimensional Kalman filter.

Our Detection Theorem, in what follows, is derived from the following identity:

LEMMA 1 (REPRODUCING GAUSSIAN DENSITIES, MALCOLM PHD THESIS) *Suppose a random vector* $\xi \in \mathbb{R}^n$, *is a normally distributed with* $\xi \sim N(\mu, \Sigma)$. *Further, suppose* A *is any matrix in* $\mathbb{M}^{m\times n}$, y *is a vector in* $\mathbb{R}^n$ *and the matrix* $B \in \mathbb{M}^{m\times m}$, *is nonsingular. Recalling the definition given for the function* $\Psi(\cdot)$, *and writing* $p(\xi)$ *for the Gaussian density function of* ξ, *then the following identity holds,*

$$\begin{aligned}\int_{\mathbb{R}^n} \Psi\big(B^{-1}(y - A\xi)\big)p(\xi)d\xi &= E\big[\exp\big\{-\tfrac{1}{2}(y - A\xi)' \operatorname{inv}(BB')(y - A\xi)\big\}\big] \\ &= (2\pi)^{-n/2}|B||BB' + A\Sigma A'|^{-\frac{1}{2}}\times \\ &\qquad \exp\big\{-\tfrac{1}{2}(y - A\mu)'(BB' + A\Sigma A')^{-1}(y - A\mu)\big\}. \end{aligned} \tag{57}$$

THEOREM 4 *The unnormalised probability* q^j_{k+1}*, corresponding to the hypothesis* H_j*, is given by the recursion*

$$q^j_{k+1} = \frac{1}{\phi(y_{k+1})\sqrt{|D_j^2 + C_j \Sigma^j_{k+1|k} C'_j|}} \times$$
$$\exp\left\{ -\tfrac{1}{2}(y_{k+1} - C_j \widehat{x}^j_{k+1|k})(D_j^2 + C_j \Sigma^j_{k+1|k} C'_j)^{-1}(y_{k+1} - C_j \widehat{x}^j_{k+1|k}) \right\} q^j_k, \quad (58)$$

where

$$\widehat{x}^j_{k+1|k} = E[x_{k+1} \mid \boldsymbol{\alpha} = \boldsymbol{f}_j, \mathcal{Y}_k], \quad (59)$$
$$\Sigma^j_{k+1|k} = E[(x_{k+1} - \widehat{x}^j_{k+1|k})\,(x_{k+1} - \widehat{x}^j_{k+1|k})' \mid \boldsymbol{\alpha} = \boldsymbol{f}_j, \mathcal{Y}_k]. \quad (60)$$

The Kalman filter, constructed under the hypothesis H_j*, provides recursive up-dates for* $\widehat{x}^j_{k+1|k}$ *and* $\Sigma^j_{k+1|k}$.

- In our example we consider a model with parameter values $e = 0.2$, $a = 0.8$ $b = 0.6$, and $d = 0.2$. We take our sampling frequency as once per day, so here $\Delta_t = 1/365$.
- A single realisation of 100 samples in length was generated by the observation model. This single realisation was subsequently used as inputs to the M individual Kalman filters and to the M detectors.
- We suppose there are 9 model hypotheses, writing $H_j = \{e_j, a_j, b_j, d_j\}$, our 9 candidate models are as follows:

$$H_1 \triangleq \{0.2, 0.1, 0.6, 0.2\}, \quad (61)$$

$$H_2 \triangleq \{0.2, 0.2, 0.6, 0.2\}, \quad (62)$$

$$H_3 \triangleq \{0.2, 0.3, 0.6, 0.2\}, \quad (63)$$

$$H_4 \triangleq \{0.2, 0.4, 0.6, 0.2\}, \quad (64)$$

$$H_5 \triangleq \{0.2, 0.5, 0.6, 0.2\}, \quad (65)$$

$$H_6 \triangleq \{0.2, 0.6, 0.6, 0.2\}, \quad (66)$$

$$H_7 \triangleq \{0.2, 0.7, 0.6, 0.2\}, \quad (67)$$

$$H_8 \triangleq \{0.2, 0.8, 0.6, 0.2\}, \quad (68)$$

$$H_9 \triangleq \{0.2, 0.9, 0.6, 0.2\}. \quad (69)$$

- Our single realisation of 100 samples in length was then separately submitted to each of the 9 Kalman filters constructed from the model hypotheses above.
- The resulting estimated detector probabilities are shown in figure 1 for H_3, H_7, $H_8$and H_9.
- The model parameter set corresponding to the true model begins to emerge as the dominant hypothesis at approximately 50 samples.
- It is worth contrasting these results shown in Figure 1 with the identical scenario, but cast as an EM algorithm, <u>200 observations</u> were used and the number of iterations up to approximate convergence of the EM algorithm was <u>150</u>.

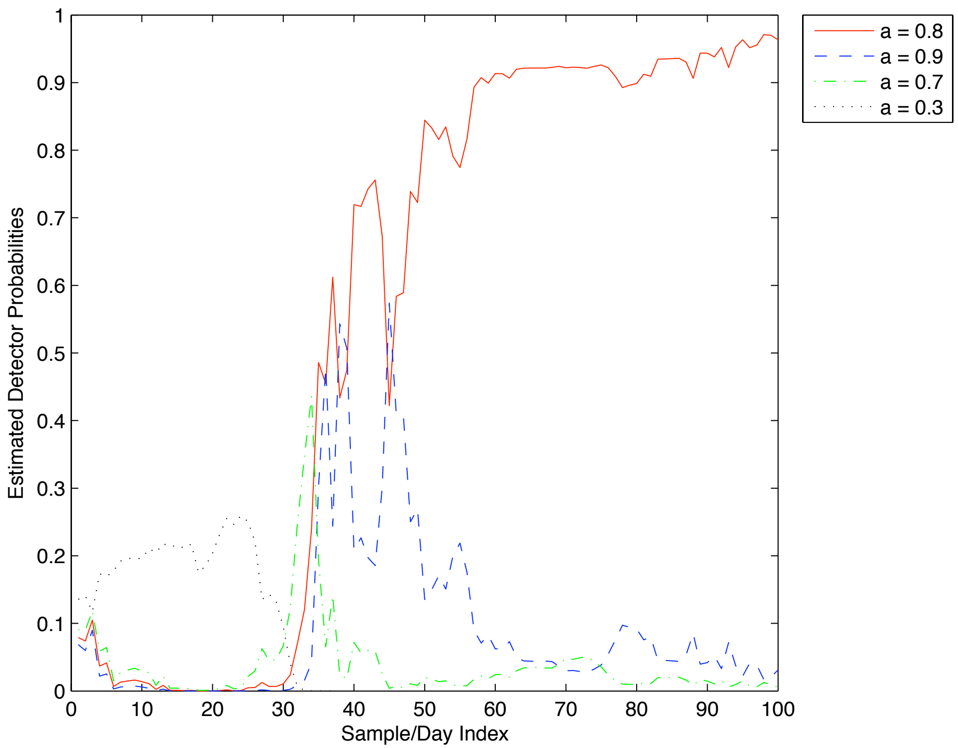

Figure 3: This plot shows estimated model hypotheses, for the particular hypotheses $H = H_3$, $H = H_7$, $H = H_8$ and $H = H_9$. The true model is H_8 with the parameter value $a = 0.8$.

Future Work

- Testing on real data sets (Pairs Trading Detection)
- The question of choosing the candidate hypotheses, that is, applying a suitable metric to ensure that all the candidate models are "suitably" separated in the space of models
- How much of the estimated probabilities $P(H = H_j \mid \mathcal{Y}_t)$, must one observe to make a decision with a certain specified degree of confidence ? SPRT $M > 2$?
- Saving the Polar Bear and reducing the cost of Oil