11 Application of Extended Kalman Filter to SPD Estimation

Zdeněk Hlávka and Marek Svojik

The state price density (SPD) carries important information on the behavior and expectations of the market and it often serves as a base for option pricing and hedging. Many commonly used SPD estimation technique are based on the observation (Breeden and Litzenberger, 1978) that the SPD $f(.)$ may be expressed as

$$f(K) = \exp\{r(T-t)\}\frac{\partial^2 C_t(K,T)}{\partial K^2}, \tag{11.1}$$

where $C_t(K,T)$ is a price of European call option with strike price K at time t expiring at time T and r denotes the risk free interest rate. An overview of estimation techniques is given in Jackwerth (1999). Kernel smoothers were in this framework applied by A¨it-Sahalia and Lo (1998), A¨it-Sahalia and Lo (2000), or Huynh, Kervella, and Zheng (2002). Some modifications of the nonparametric smoother allowing to apply no-arbitrage constraints were proposed, e.g., by A¨it-Sahalia and Duarte (2003), Bondarenko (2003), or Yatchew and Härdle (2006). Apart of the choice of a suitable estimation method, Härdle and Hlávka (2005) show that the covariance structure of the observed option prices carries additional important information that should to be considered in the estimation procedure. Härdle and Hlávka (2005) suggest a simple and easily applicable approximation of the covariance. A more detailed discussion of option price errors may be found in Renault (1997).

In this chapter, we will estimate the SPD from observed call option prices using the well-known Kalman filter, invented already in the early sixties and marked by Harvey (1989). Kalman filter may be shortly described as a statistical method used for estimation of the non-observable component of a state-space model and it already became an important econometric tool for financial and economic estimation problems in continuous time finance. More precisely, the Kalman filter is a recursive procedure for computing the optimal estimator of the state vector $\boldsymbol{\xi}_i$ at time i, based on information available at time i. For derivation of the Kalman filter, we focus on the general system characterized by a *measurement equation*

$$\mathbf{y}_i = \mathbf{B}_i(\boldsymbol{\psi})\boldsymbol{\xi}_i + \boldsymbol{\varepsilon}_i(\boldsymbol{\psi}), \qquad i = 1, \ldots, n, \tag{11.2}$$

and a *transition equation*

$$\boldsymbol{\xi}_i = \boldsymbol{\Phi}_i(\boldsymbol{\psi})\boldsymbol{\xi}_{i-1} + \boldsymbol{\eta}_i(\boldsymbol{\psi}), \qquad i = 1, \ldots, n, \tag{11.3}$$

where $\mathbf{y}_i$ is the g-dimensional vector of the observable variables and $\boldsymbol{\xi}_i$ denotes the unobservable k-dimensional *state vector*, with unknown parameters $\boldsymbol{\psi}$, a known matrix $\mathbf{B}_i(\boldsymbol{\psi})$, and a noise term $\boldsymbol{\varepsilon}_i(\boldsymbol{\psi})$ of serially uncorrelated disturbances with zero mean and variance matrix $\mathbf{H}_i(\boldsymbol{\psi})$. The symbols used in the transition equation (11.3) are the *transition matrix* $\boldsymbol{\Phi}_i(\boldsymbol{\psi})$ and a zero mean Gaussian noise term $\boldsymbol{\eta}_i(\boldsymbol{\psi})$ with a known variance matrix $\mathbf{Q}_i(\boldsymbol{\psi})$. The specification of the state space model is completed by assuming independence between the error terms $\boldsymbol{\varepsilon}_t(\boldsymbol{\psi})$ and $\boldsymbol{\eta}_t(\boldsymbol{\psi})$. Additionally, we assume that these error terms are uncorrelated with the normally distributed *initial state vector* $\boldsymbol{\xi}_0$ having expected value $\boldsymbol{\xi}_{0|0}$ and variance matrix $\boldsymbol{\Sigma}_{0|0}$.

The state-space model (11.2)–(11.3) is suitable for situations in which, instead of being able to observe the state vector $\boldsymbol{\xi}_i$ directly, we can only observe some noisy function $\mathbf{y}_i$ of $\boldsymbol{\xi}_i$. The problem of determining the state of the system from noisy measurements $\mathbf{y}_i$ is called *estimation*. *Filtering* is a special case of estimation with the objective of obtaining an estimate of $\boldsymbol{\xi}_i$ given observations up to time i. It can be shown that the optimal estimator of $\boldsymbol{\xi}_i$, i.e., minimizing the mean squared error (MSE), is the mean of the conditional distribution of the state vector $\boldsymbol{\xi}_i$. When estimating $\boldsymbol{\xi}_i$ using information up to time s, we denote the conditional expectation of $\boldsymbol{\xi}_i$ given $\mathcal{F}_s$ for convenience by $\boldsymbol{\xi}_{i|s} = \mathsf{E}[\boldsymbol{\xi}_i|\mathcal{F}_s]$. The conditional variance matrix of $\boldsymbol{\xi}_i$ given $\mathcal{F}_s$ is denoted as $\boldsymbol{\Sigma}_{i|s} = \mathsf{Var}[\boldsymbol{\xi}_i|\mathcal{F}_s]$.

In our case, we will see that the relationship between the state and observed variables is nonlinear and the problem has to be linearized by Taylor expansion.

11.1 Linear Model

Let us remind that the payoff for a call option is given by

$$(S_T - K)_+ = \max(S_T - K, 0).$$

Let $\mathbf{C}_t(K, T)$ be the call pricing function of a European call option with strike price K observed at time t and expiring at time T. We consider a call option with this payoff. Let S_T denotes the price of the underlying asset at T, and r the risk free interest rate. Then, the fair price $\mathbf{C}_t(K,T)$ of a European call option at the current time t may be expressed as the discounted expected value of the payoff $(S_T - K)_+$ with respect to the SPD $f(.)$, i.e.,

$$\mathbf{C}_t(K,T) = e^{-r(T-t)} \int_0^{+\infty} (S_T - K)_+ f(S_T) dS_T. \tag{11.4}$$

Clearly, the call pricing function $\mathbf{C}_t(K,T)$ is monotone decreasing and convex in K.

In the rest of this chapter, we will assume that the discount factor $e^{-r(T-t)}$ in (11.4) is equal to 1. In practice, this may be easily achieved by dividing the observed option prices by this known discount factor.

In (11.1), we have already seen that the SPD may be expressed as the discounted second derivative of the call pricing function $\mathbf{C}_t(K,T)$ with respect to the strike price K. We will use this relationship to construct an SPD estimator based on the observed call option prices.

11.1.1 Linear Model for Call Option Prices

On a fixed day t, the i-th observed option price corresponding to the time of expiry T will be denoted as $\mathbf{C}_i = \mathbf{C}_{t,i}(\mathbf{K}_i, T)$, where K_i denotes the corresponding strike price. The vector of all observed option prices will be denoted as $\boldsymbol{\mathcal{C}} = (\mathbf{C}_1, \ldots, \mathbf{C}_n)^\top$. Without loss of generality, we assume that the corresponding vector of the strike prices $\mathcal{K} = (K_1, \ldots, K_n)^\top$ has the following structure

$$\boldsymbol{\mathcal{K}} = \begin{pmatrix} \mathbf{K}_1 \\ \mathbf{K}_2 \\ \vdots \\ \mathbf{K}_n \end{pmatrix} = \begin{pmatrix} k_1 1_{n_1} \\ k_2 1_{n_2} \\ \vdots \\ k_n 1_{n_p} \end{pmatrix},$$

where $k_1 < k_2 < \cdots < k_p$ are the p distinct values of the strike prices, 1_{n_j} denotes a vector of ones of length n_j, and $n_j = \sum_{i=1}^{n} \mathbf{1}(\mathbf{K}_i = k_j)$.

The further assumptions and constraints that have to be satisfied by the developing the linear model are largely taken from Härdle and Hlávka (2005). We impose only constraints that guarantee that the estimated function is probability density, i.e., it is positive and it integrates to one. The SPD is parameterized by assuming that for a fixed day t and time to maturity $\tau = T - t$, the i-th observed option price $\mathbf{C}_i$ corresponding to strike price $\mathbf{K}_i$, the option prices $\mathbf{C}_i = \mathbf{C}_{t,i}(\mathbf{K}_i, T)$ follows the linear model

$$\mathbf{C}_{t,i}(\mathbf{K}_i, T) = \mu(\mathbf{K}_i) + \epsilon_i, \tag{11.5}$$

where $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_n)^\top \sim N(\mathbf{0}, \Sigma)$ is random vector of correlated normally distributed random errors.

In the next section, we will parameterize the vector of the mean option prices $\mu(.)$ in terms of the state price density. This parameterization will allow us to derive SPD estimators directly from the linear model (11.5).

11.1.2 Estimation of State Price Density

In Härdle and Hlávka (2005), it was suggested to rewrite the vector of the conditional means $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_p)^\top$ in terms of the parameters $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_{p-1})^\top$ as

$$\boldsymbol{\mu} = \boldsymbol{\Delta\beta}, \tag{11.6}$$

where

$$\boldsymbol{\Delta} = \begin{pmatrix} 1 & \Delta_p^1 & \Delta_{p-1}^1 & \Delta_{p-2}^1 & \cdots & \Delta_3^1 & \Delta_2^1 \\ 1 & \Delta_p^2 & \Delta_{p-1}^2 & \Delta_{p-2}^2 & \cdots & \Delta_3^2 & 0 \\ \vdots & & & & & & \vdots \\ 1 & \Delta_p^{p-2} & \Delta_{p-1}^{p-2} & 0 & \cdots & 0 & 0 \\ 1 & \Delta_p^{p-1} & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \tag{11.7}$$

and $\Delta_j^i = \max(k_j - k_i, 0)$ denotes the positive part of the distance between k_i and k_j, i.e. the i-th and the j-th $(1 \le i \le j \le p)$ sorted distinct observed values of the strike price.

The vector of parameters β in (11.6) may be interpreted as an estimate of the second derivatives of the call pricing function and consequently, according to (11.1), also as the estimator of the state price density.

The constraints on the conditional means μ_j such as positivity, monotonicity and convexity can be reexpressed in terms of β_j—it suffices to request that $\beta_j > 0$ for $j = 0, \dots, p-1$ and that $\sum_{j=2}^{p-1} \beta_j \le 1$.

Using this notation, the linear model for the observed option prices $\mathbf{C}$ is obtained by

$$\mathbf{C}(\boldsymbol{\mathcal{K}}) = \boldsymbol{\mathcal{X}}_{\boldsymbol{\Delta}}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \tag{11.8}$$

where $\boldsymbol{\mathcal{X}}_{\boldsymbol{\Delta}}$ is the design matrix obtained by repeating each row of the matrix $\boldsymbol{\Delta}$ n_i-times for $i = 1, \dots, p$.

11.1.3 State-Space Model for Call Option Prices

In order to apply Kalman filter without any constraints on the resulting SPD estimates (β_i, $i = 0, \dots, p-1$), we rewrite the linear model (11.8) in a state-space form for the i-th observation on a fixed day t:

$$\mathbf{C}_i(\boldsymbol{\mathcal{K}}) = \boldsymbol{\mathcal{X}}_{\boldsymbol{\Delta}}\boldsymbol{\beta}_i + \boldsymbol{\varepsilon}_i, \tag{11.9}$$

$$\boldsymbol{\beta}_i = \boldsymbol{\beta}_{i-1} + \boldsymbol{\eta}_i, \tag{11.10}$$

where $\boldsymbol{\mathcal{X}}_{\boldsymbol{\Delta}}$ is the design matrix from (11.8) and $\boldsymbol{\varepsilon}_i \sim \mathrm{N}\big(\mathbf{0}, \sigma^2\mathbf{I}\big)$ and $\boldsymbol{\eta}_i \sim \mathrm{N}\big(\mathbf{0}, \nu^2\delta_i\mathbf{I}\big)$ are uncorrelated random vectors. We assume that the variance of η_i depends linearly on the time δ_i between the i-th and the $(i-1)$-st trade.

In the following, we determine the Kalman filter in a standard way. The standard approach has to be only slightly modified as in every step i we observe only option price $C_i(K_i)$ corresponding to only one strike price K_i.

Prediction step In the prediction step, we forecast the state vector by calculating the conditional moments of the state variables given the information up to time $t-1$ to obtain the *prediction equations*

$$\boldsymbol{\beta}_{i|i-1} = \mathsf{E}(\boldsymbol{\beta}_i|\mathcal{F}_{i-1}) = \boldsymbol{\beta}_{i-1|i-1}, \tag{11.11}$$

$$\boldsymbol{\Sigma}_{i|i-1} = \boldsymbol{\Sigma}_{i-1|i-1} + \nu^2\delta_i\mathbf{I}. \tag{11.12}$$

Updating step Denoting by $\boldsymbol{\Delta}_i$ the i-th row of the design matrix $\boldsymbol{\mathcal{X}}_{\boldsymbol{\Delta}}$, i.e., the row corresponding to the i-th observed strike price K_i, we arrive to the *updating equations*

$$\boldsymbol{\beta}_{i|i} = \boldsymbol{\beta}_{i|i-1} + \mathbf{K}_i I_i, \tag{11.13}$$

$$\boldsymbol{\Sigma}_{i|i} = (\mathbf{I} - \mathbf{K}_i\boldsymbol{\Delta}_i)\,\boldsymbol{\Sigma}_{i|i-1}, \tag{11.14}$$

where

$$I_i = C_i(K_i) - C_{i|i-1}(K_i) = C_i(K_i) - \boldsymbol{\Delta}_i\boldsymbol{\beta}_{i|i-1}$$

is the *prediction error* with variance $F_{i|i-1} = \mathsf{Var}(I_i|\mathcal{F}_{i-1}) = \sigma^2 + \boldsymbol{\Delta}_i\boldsymbol{\Sigma}_{i|i-1}\boldsymbol{\Delta}_i^\top$ and $\mathbf{K}_i = \boldsymbol{\Sigma}_{i|i-1}\boldsymbol{\Delta}_i^\top F_{i|i-1}^{-1}$ is the *Kalman gain*.

The prediction and updating equations (11.11)–(11.14) jointly constitute the linear Kalman filter for a European call option. Unfortunately, in this case the practical usefulness of the linear Kalman filter is limited as the resulting SPD estimator does not have to be probability density. A more realistic nonlinear model is presented in the following Section 11.2.

11.2 Extended Kalman Filter and Call Options

In the following, we constrain the vector of parameters $\boldsymbol{\beta}_i = (\beta_0, \dots, \beta_{p-1})^\top$ so that it may always be reasonably interpreted as a probability density. We propose a reparameterization of the model in terms of parameters $\boldsymbol{\xi}_i = (\xi_0, \dots, \xi_{p-1})^\top$ via a smooth function $\mathbf{g}_i(\cdot) = \big(g_0(\cdot), \dots, g_{p-1}(\cdot)\big)^\top$ by setting

$$\beta_0 = g_0(\boldsymbol{\xi}_i) = \exp(\xi_0), \tag{11.15}$$

$$\beta_k = g_k(\boldsymbol{\xi}_i) = S^{-1}\exp(\xi_k), \quad \text{for } k = 1, \dots, p-1, \tag{11.16}$$

where $S = \sum_{j=1}^{p-1} \exp(\xi_j)$ simplifies the notation.

Obviously, $\sum_{j=1}^{p-1} \beta_j = 1$ and $\beta_j > 0$, $j = 0, \dots, p-1$. This means that the parameters β_j, $j = 1, \dots, p-1$ are positive and integrate to one and may be interpreted as a reasonable estimates of the values of the SPD.

The linear model for the option prices (11.8) rewritten in terms of $\boldsymbol{\xi}_i$ leads a nonlinear state space model given by the measurement equation

$$\mathbf{C}_i(\boldsymbol{\mathcal{K}}) = \boldsymbol{\mathcal{X}}_{\boldsymbol{\Delta}} \mathbf{g}_i(\boldsymbol{\xi}_i) + \boldsymbol{\varepsilon}_i, \tag{11.17}$$

and the transition equation

$$\boldsymbol{\xi}_i = \boldsymbol{\xi}_{i-1} + \boldsymbol{\eta}_i, \tag{11.18}$$

where $\boldsymbol{\varepsilon}_i \sim \mathrm{N}\big(\mathbf{0}, \sigma^2 \mathbf{I}\big)$ and $\boldsymbol{\eta}_i \sim \mathrm{N}\big(\mathbf{0}, \nu^2 \delta_i \mathbf{I}\big)$ satisfy the same assumptions as in Section 11.1.3.

The extended Kalman filter for the above problem may be linearized by Taylor expansion using the Jacobian matrix $\mathbf{B}_{i|i-1}$ computed in $\boldsymbol{\xi}_i = \boldsymbol{\xi}_{i|i-1}$:

$$\begin{aligned} \mathbf{B}_{i|i-1} &= \left.\frac{\partial \mathbf{g}_i(\boldsymbol{\xi}_i)}{\partial \boldsymbol{\xi}_i^\top}\right|_{\boldsymbol{\xi}_i = \boldsymbol{\xi}_{i|i-1}} \qquad (11.19) \\ &= \frac{1}{S^2} \begin{pmatrix} S^2 e^{\xi_0} & 0 & 0 & \cdots & 0 \\ 0 & e^{\xi_1}\big(S - e^{\xi_1}\big) & -e^{\xi_1+\xi_2} & \cdots & -e^{\xi_1+\xi_{p-1}} \\ 0 & -e^{\xi_2+\xi_1} & e^{\xi_2}\big(S - e^{\xi_2}\big) & \cdots & -e^{\xi_2+\xi_{p-1}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & -e^{\xi_{p-1}+\xi_1} & -e^{\xi_{p-1}+\xi_2} & \cdots & e^{\xi_{p-1}}\big(S - e^{\xi_{p-1}}\big) \end{pmatrix}. \end{aligned}$$

Now, the linearized version of the Kalman filter algorithm for model (11.17)–(11.18) is straightforward. Similarly as in Section 11.1.3, we obtain *extended prediction equations*

$$\begin{aligned} \boldsymbol{\xi}_{i|i-1} &= \boldsymbol{\xi}_{i-1|i-1}, && (11.20) \\ \boldsymbol{\Sigma}_{i|i-1} &= \boldsymbol{\Sigma}_{i-1|i-1} + \nu^2 \delta_i \mathbf{I}, && (11.21) \end{aligned}$$

and *extended updating equations*

$$\begin{aligned} \boldsymbol{\xi}_{i|i} &= \boldsymbol{\xi}_{i|i-1} + \mathbf{K}_i I_i, && (11.22) \\ \boldsymbol{\Sigma}_{i|i} &= \left(\mathbf{I} - \mathbf{K}_i \boldsymbol{\Delta}_i \mathbf{B}_{i|i-1}\right) \boldsymbol{\Sigma}_{i|i-1}, && (11.23) \end{aligned}$$

where $I_i = C_i(K_i) - \boldsymbol{\Delta}_i \mathbf{g}_i(\boldsymbol{\xi}_{i|i-1})$ is the prediction error, $F_{i|i-1} = \mathsf{Var}(I_i|\mathcal{F}_{i-1}) = \sigma^2 + \boldsymbol{\Delta}_i \mathbf{B}_{i|i-1} \boldsymbol{\Sigma}_{i|i-1} \mathbf{B}_{i|i-1}^\top \boldsymbol{\Delta}_i^\top$ its variance, and $\mathbf{K}_i = \boldsymbol{\Sigma}_{i|i-1} \mathbf{B}_{i|i-1}^\top \boldsymbol{\Delta}_i^\top F_{i|i-1}^{-1}$ the Kalman gain.

The recursive equations (11.20)–(11.23) form the extended Kalman filter recursions and lead the vector $\mathbf{g}_i(\boldsymbol{\xi}_i) = \boldsymbol{\beta}_i$ representing estimates of the SPD.

11.3 Empirical Results

In this section, the extended Kalman filter is used to estimate SPD from DAX call option prices. In other words, our objective is to estimate the call function $\mathbf{C}_t(K, T)$ subject to monotonicity and convexity constraints, i.e., the constraint that the implied SPD is non-negative and it integrates to one.

We choose data over a sufficiently brief time span so that the time to maturity τ, the interest rate r, and both the current time t and the time of expiry T may be considered as constant. The full data set contains observed call and put option prices for various strike prices and maturities τ. From now on, for each trading day, we consider only a subset containing the call options $\mathbf{C}_{t,i}(K_i, T)$, $i = 1, \ldots, n$ with the shortest time to expiry $\tau = T - t$. In 1995, we have few hundreds such observations each day. In 2003, the number of daily observations increases to thousands.

Apart of the strike prices K_i and option prices $\mathbf{C}_{t,i}(K_i, T)$, the data set contains also information on the risk-free interest rate r, the time of trade (given in seconds after midnight), the current value of the underlying asset (DAX), time to expiry, and type of the option (Call/Put).

As the risk-free interest rate r and the time to expiry $T - t$ are known and given in our data set, we may work with option prices corrected by the known discount factor $e^{-r(T-t)}$. This modification guarantees that the second derivative of the discounted call pricing function is equal to the state price density.

11.3.1 Extended Kalman Filtering in Practice

In order to implement the Kalman filter in practice, we need to:

1. set the *initial values* of unknown parameters,
2. *estimate the unknown parameters* from data.

Initialization In order to use the extended version Kalman filter, we have to choose initial values $\mathbf{\Sigma}_{0|0}$ and $\boldsymbol{\beta}_{0|0}$ and variances of both error terms $\boldsymbol{\varepsilon}_i$ and $\boldsymbol{\eta}_i$. We choose initial $\mathbf{\Sigma}_{0|0} = \mathbf{I}$ and

$$\boldsymbol{\beta}_{0|0} = \left(\underbrace{\widehat{\mathsf{E}}\{C(k_p)\}}_{\beta_0}, \underbrace{\frac{1}{p-1}, \dots, \frac{1}{p-1}}_{p-1} \right),$$

i.e., β_0 is set as the sample mean of option prices corresponding to the largest strike price k_p. The remaining values, defining the initial distribution of the SPD, are set uniformly.

The parameter σ^2 may be interpreted as the standard error of the option price in Euros. The interpretation of the parameter ν^2 is more difficult and it depends on the time intervals between consecutive trades and on the range of the observed strike prices. For the first run of the algorithm, we set the variance matrices as $\mathsf{Var}[\boldsymbol{\varepsilon}_i] = \sigma^2\mathbf{I}$ and $\mathsf{Var}[\boldsymbol{\eta}_i] = \nu^2\delta_i\mathbf{I}$, with $\sigma^2 = 1$ and

$$\nu^2 = 1/[\{(\max_{i=1,\dots,n} K_i - \min_{i=1,\dots,n} K_i)/2\}^2 \min_{i=1,\dots,n} \delta_i].$$

This choice is quite arbitrary but it reflects that the parameter σ^2 should be small (in Euros) and that the parameter ν^2 is related to the time and to the range of the observed strike prices. Note that these are only initial values and more realistic estimates are obtained in the next iterations of the extended Kalman filter.

Extended Kalman filter Given the starting values $\boldsymbol{\beta}_{0|0}$, $\mathbf{\Sigma}_{0|0}$, σ^2, and ν^2, the extended Kalman filter is given by equations (11.20)–(11.23). The nonlinear projections $\mathbf{g}_i(.)$ guarantee that the state vector $\boldsymbol{\beta}_i = \mathbf{g}_t(\boldsymbol{\xi}_i)$ satisfies the required constraints.

Parameter estimation The unknown parameters σ^2 and ν^2 are estimated by Maximum Likelihood (ML) method. More precisely, we use the *prediction error decomposition of the likelihood function* described in Kellerhals (2001, Chapter 5), the resulting log-likelihood is then maximized numerically. Note that another approach to parameter estimation based on the Kalman smoother and EM-algorithm is described in Harvey (1989, Section 4.2.4).

The behavior of the extended Kalman filter depends also on the choice of the starting value $\boldsymbol{\beta}_{0|0}$. Assuming that the shape of the SPD doesn't change too much during the day, we may improve on the initial "uniform" SPD by taking $\boldsymbol{\beta}_{n|n}$, shift the corresponding SPD by the difference of the value of the underlying asset, and by using the resulting set of parameters as the starting value $\boldsymbol{\beta}_{0|0}$. In practice, one might use the final estimator $\beta_{n|n}$ from day t as the initial estimator on the next day $t+1$

Kalman filter iterations Combining the initial parameter choice, the Kalman filter, and the parameter estimation, we obtain the following iterative algorithm:

1. Choose the initial values.
2. Run the extended Kalman filter (11.20)–(11.23) with current values of the parameters.
3. Use the Kalman filter predictions to estimate the parameters σ^2 and ν^2 by numerical maximization of the log-likelihood and update the initial values $\boldsymbol{\beta}_{0|0}$ and $\boldsymbol{\Sigma}_{0|0}$ using $\boldsymbol{\beta}_{n|n}$ and $\boldsymbol{\Sigma}_{n|n}$.
4. Either stop the algorithm or return to step 2 depending on the chosen stopping rule.

In practice, the stopping rule for the above iterative algorithm may be based on the values of the log-likelihood obtained in step 3 of the iterative algorithm. In the following real life examples, we will run fixed number of iterations as an illustration.

11.3.2 SPD Estimation in 1995

The first example is using data from two trading days in 1995; these are the two data sets as in Härdle and Hlávka (2005). The call option prices observed on 11th (January 14th) and 12th trading day (January 15th) in 1995 are plotted on the left-hand side graphics in Figures 11.1 and 11.2. The main difference between these two trading days is that the strike prices traded on 15th January cover larger range of strike prices. This means that

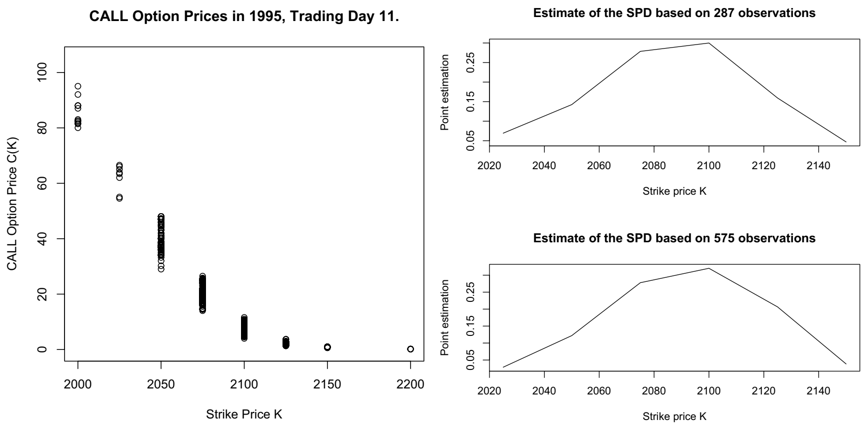

Figure 11.1. European call option prices with shortest time to expiry plotted against strike price K (left) and two of the filtered SPD estimates (right) on JAN-14-1995. XFGKF1995a

also the support of the estimated SPDs will be larger on January 15th than on January 14th.

JAN-14-1995 Using the data from January 14th, 1995, we ran 10 iterations of the algorithm described in Section 11.3.1. The resulting parameter estimates, $\widehat{\sigma}^2 = 0.0111$ and $\widehat{\nu}^2 = 2.6399$, seem to be stable. In the last four iterations, estimates of σ^2 vary between 0.0081 and 0.0115 and estimates of ν^2 are varying between 2.4849 and 2.6399.

The Kalman filter provides SPD estimate in each time $i = 1, \dots, n$ and we thus obtain altogether $n = 575$ estimates of $\boldsymbol{\beta}_i$ during this one day. Two of these filtered SPD estimates on JAN-14-1995 are displayed on the on the right-hand side of Figure 11.1; the upper plot shows the estimator at time $i_1 = 287 \doteq n/2$ (12:45:44.46) and the lower plot the estimator at time $i_2 = n = 575$ (15:59:59.78), i.e., The lower plot contains the estimator of the SPD at the end of this trading day.

Both estimates look very similar but the latter one is shifted a bit towards higher values. This shift is clearly due to a change in the value of the underlying asset (DAX) from 2087.691 to 2090.17 during the same time period.

JAN-15-1995 Next, the same technique is applied to data observed on January 15th, 1995, see Figure 11.2. Two of the resulting filtered SPD estimates

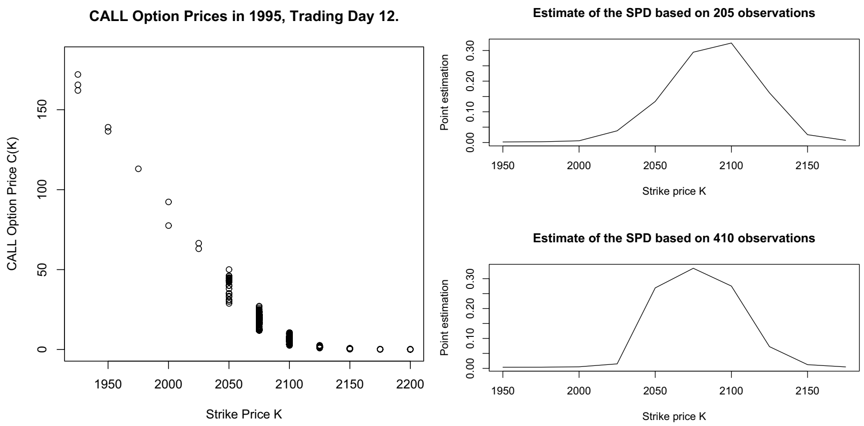

Figure 11.2. European call option prices with the shortest time to expiry plotted against strike price K (left) and two of the filtered SPD estimates (right) on JAN-15-1995. XFGKF1995b

are plotted in the graphics on the right-hand side of Figure 11.2. The SPD estimator calculated at the time $i_1 = n/2 = 205$ (12:09:14.65) is almost identical to the final estimate from January 14th; the most visible difference is the larger support for the estimated SPD on JAN-15-1995. At the end of this trading day, for $i_2 = n = 410$ (15:59:52.14), the estimate is shifted a bit to the left and more concentrated. The shift to the left corresponds again to a decrease in the value of the DAX from 2089.377 to 2075.989.

The parameter estimates obtained after 10 iterations of the algorithm described in Section 11.3.1 are $\widehat{\sigma}^2 = 0.0496$ and $\widehat{\nu}^2 = 0.621$. Smaller value of $\widehat{\nu}^2$ seemingly suggests that the SPD was changing more slowly on JAN-15-1995 but this parameter must be interpreted with a caution as its scale depends also on the size of the time interval between consecutive trades and on the range of the observed strike prices.

11.3.3 SPD Estimation in 2003

The next example is using the most recent data set in our database. On February 25th, 2003, we observe altogether 1464 call option prices with the shortest time to expiry. Compared to the situation in 1995, the option markets in 2003 are more liquid and the number of distinct strike prices included in the data set is larger than in 1995. Our data set contains 30 distinct strike prices on FEB-25-2003 compared to 8 on JAN-14-1995 and 12 on JAN-15-1995.

The call option prices observed on FEB-25-2003 are plotted as a function of their strike price on the left-hand side plot in Figure 11.3.

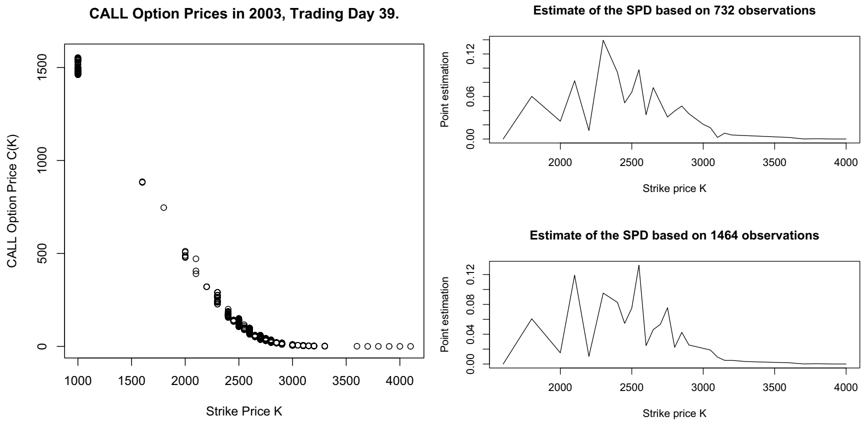

Figure 11.3. European call option prices with shortest time to expiry plotted against strike price K on FEB-25-2003, $n = 1464$ observed prices (left) and the resulting SPD estimates after 10 iterations (right). XFGKF2003

After ten iterations of the iterative extended Kalman filtering algorithm described in Section 11.3.1, we obtain parameter estimates $\widehat{\sigma}^2 = 0.0324$ and $\widehat{\nu}^2 = 3.1953$. The corresponding SPD estimates for times $i_1 = n/2 = 732$ and $i_2 = n = 1464$ are plotted on the right-hand side of Figure 11.3.

On FEB-25-2003, the resulting estimates do not look very much like a typical (smooth and unimodal) probability densities. Instead, we observe a lot of spikes and valleys. This is due to the fact that the algorithm does not penalize non-smoothness and the reparameterization (11.15)–(11.16) guarantees only that the resulting SPD estimates are positive and integrate to one.

In order to obtain more easily interpretable results, the resulting estimates may be smoothed using, e.g., the Nadaraya-Watson kernel regression estimator (Nadaraya, 1964; Watson, 1964). As the smoothing of the vector $\boldsymbol{\beta}_{n|n}$ corresponds to a multiplication with a (smoothing) matrix, say $\boldsymbol{S}$, the smoothing step may be implemented after the Kalman filtering, see Härdle (1991) or Simonoff (1996) for more details on kernel regression.

Using the variance matrix $\boldsymbol{\Sigma}_{n|n}$ from the filtering step of the extended Kalman filtering algorithm, we calculate the variance matrix of ${\boldsymbol{\xi}^{smooth}}_{n|n} = \boldsymbol{S}\boldsymbol{\xi}_{n|n}$ as $\mathsf{Var}{\boldsymbol{\xi}^{smooth}}_{n|n} = \boldsymbol{S}\boldsymbol{\Sigma}_{n|n}\boldsymbol{S}^\top$. This leads an approximation of the asymptotic variance of the smooth SPD estimate ${\boldsymbol{\beta}^{smooth}}_{n|n} = \boldsymbol{g_n}({\boldsymbol{\xi}^{smooth}}_{n|n})$ as $\mathsf{Var}{\boldsymbol{\beta}^{smooth}}_{n|n} = \boldsymbol{B}_n\boldsymbol{S}\boldsymbol{\Sigma}_{n|n}\boldsymbol{S}^\top\boldsymbol{B}_n^\top$, where $\boldsymbol{B}_n$ now denotes the Jacobian matrix (11.19) calculated in ${\boldsymbol{\xi}^{smooth}}_{n|n}$.

The resulting smooth SPD estimate at the end of the trading day (time n) ${\boldsymbol{\beta}^{smooth}}_{n|n}$ with pointwise asymptotic 95% confidence intervals obtained as ${\boldsymbol{\beta}^{smooth}}_{n|n} \pm 1.96\sqrt{\mathrm{diag}(\mathsf{Var}{\boldsymbol{\beta}^{smooth}}_{n|n})}$ is plotted in Figure 11.4.

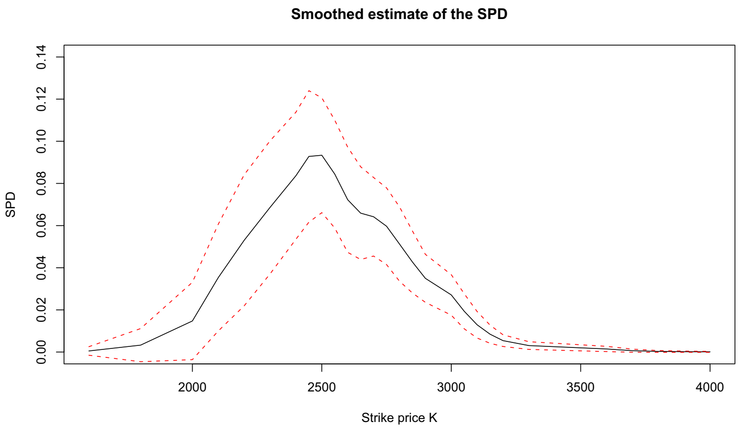

Figure 11.4. Smoothed SPD estimate on FEB-25-2003, $n = 1464$ with pointwise asymptotic confidence intervals. XFGKF2003

11.4 Conclusions

We presented and illustrated the application of extended Kalman filtering towards arbitrage free SPD estimation.

An application of the extended Kalman filtering methodology on real-world data sets in Section 11.3 shows that this method provides very good results for data sets with small number of distinct strike prices, see Figures 11.1 and 11.2.

When the number of distinct strike price increases, the linear model becomes overparameterized, and the resulting SPD estimators are not smooth anymore, see Figure 11.3. However, even in this case, the SPD estimator captures quite well the general shape of the SPD and smooth SPD estimator may be obtained by applying, e.g., the Nadaraya-Watson kernel regression estimator allowing also easy calculation of pointwise asymptotic confidence intervals.

Compared to other commonly used estimation techniques, the extended Kalman filtering methodology is able to capture the intra-day development of the SPD and it allows to update the estimates dynamically whenever new information becomes available. The extended Kalman filtering methodology combined with kernel smoothing is fast, easily applicable, and it provides interesting insights.

Bibliography

A¨**i**t-Sahalia, Y. and Duarte, J., 2003, Nonparametric option pricing under shape restrictions, *Journal of Econometrics* 116, 9–47.

A¨**i**t-Sahalia, Y. and Lo, A.W., 1998, Nonparametric estimation of state-price densities implicit in financial asset prices, *Journal of Finance* 53, 499–547.

A¨**i**t-Sahalia, Y. and Lo, A.W., 2000, Nonparametric risk management and implied risk aversion, *Journal of Econometrics* 94, 9–51.

Bondarenko, O., 2003, Estimation of risk-neutral densities using positive convolution approximation, *Journal of Econometrics* 116, 85–112.

Breeden, D. and Litzenberger, R., 1978, Prices of State-Contingent Claims Implicit in Option Prices, *Journal of Business* 51, 621–651.

Franke, J., Härdle, W. and Hafner, Ch., 2008, Statistics of Financial Markets. Springer, Berlin.

Härdle, W., 1991, Applied Nonparametric Regression. Cambridge University Press, Cambridge.

Härdle, W. and Hlávka, Z., 2005, Dynamics of State Price Densities. *Sonderforschungsbereich 649 Discussion Paper 2005-021, Humboldt-Universität zu Berlin.*

Harvey, A. C., 1989, Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge University Press, Cambridge.

Huynh, K., Kervella, P., and Zheng, J., 2002, Estimating state-price densities with nonparametric regression. In Härdle, Kleinow and Stahl, eds., Applied quantitative finance, Springer Verlag, Heidelberg 171–196.

Jackwerth, J.C., 1999, Option-implied risk-neutral distributions and implied binomial trees: a literature review, *Journal of Derivatives* 7, 66–82.

Kellerhals, B. P., 2001, Financial Pricing Models in Continuous Time and Kalman Filtering. Springer, Heidelberg.

Nadaraya, E. A., 1964, On estimating regression, *Theory of Probability and its Applications* 9 (1), 141–142.

Renault, E., 1997, Econometric models of option pricing errors. In Kreps and Wallis, eds., Advances in Economics and Econometrics: Theory and Applications, Seventh World Congress, Volume III, Cambridge University Press, Cambridge, 223–278.

Simonoff, J. S., 1996, Smoothing Methods in Statistics. Springer, New York.

Svoj´ik, M., 2007, Application of Kalman Filtering, Diploma Thesis, Charles University in Prague, Faculty of Mathematics and Physics.

Watson, G. S., 1964, Smooth regression analysis, *Sankhya, Ser. A* 26, 359–372.

Yatchew, A. and Härdle, W., 2006, Nonparametric state price density estimation using constrained least squares and the bootstrap, *Journal of Econometrics* 133(2), 579-599.