Confidence Intervals for State Price Densities

Zdeněk Hlávka

September 5, 2003

Abstract

The state price density is a second derivative of the discounted European options prices with respect to the strike price. We use Maximum Likelihood method to derive a simple estimator of the curve such that it is decreasing, convex and its second derivative integrates to one. Confidence intervals for this estimator can be constructed using standard Maximum Likelihood theory. The method works well in praxis as illustrated on the DAX option prices data.

Key words and Phrases: option pricing, state price density estimation, nonlinear least squares, confidence intervals

1 Introduction

The fair price of European option with payoff $(S_T - K)_+ = \max(S_T - K, 0)$, with S_T denoting the price of the stock at time T, K the strike price, and r the risk free interest rate, can be written as

$$C_t(K,T) = \exp\{-r(T-t)\} \int_0^{+\infty} (S_T - K)_+ f(S_T) dS_T, \tag{1}$$

i.e., as the discounted expected value of the payoff with respect to the so-called state price density $f(S_T)$. The state price density (SPD) is widely acknowledged to bear important information on the behaviour and expectations of the market. An important application of SPD is that it allows to price options with complicated payoff functions simply by (numerical) integration of the payoff with respect to this density.

Prices $C_t(K,T)$ of European options with strike price K observed at time t and expiring at time T allow to deduce the state price density in the following form (Breeden and Litzenberger 1978)

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

Equation (2) is often used to estimate the state price density by the means of nonparametric regression. Kernel smoothers were in this framework proposed and succesfully applied by, e.g., Aït-Sahalia and Lo (1998) or Aït-Sahalia, Wang and Yared (2000). Another, more sophisticated approach based on nonparametric least squares which allows to include the required constraints is described and applied on simulated data in Härdle and Yatchew (2001).

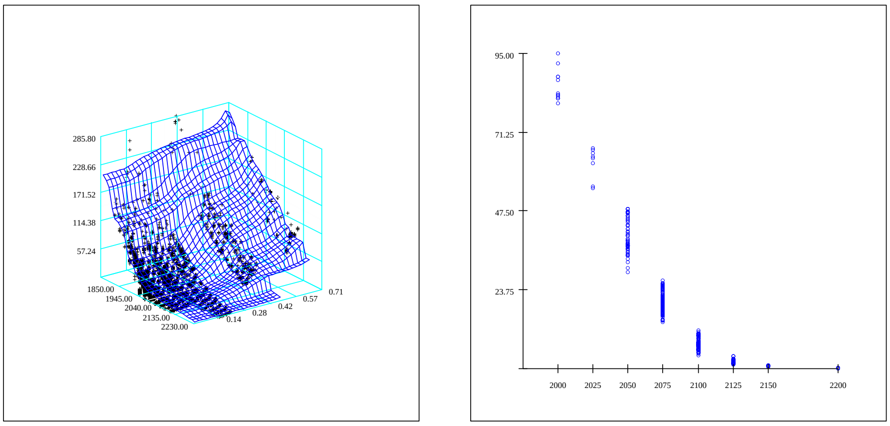

Figure 1: Option prices plotted against strike price and time to maturity with two-dimensional kernel regression surface (left) and scatterplot of the option prices against strike price (right) on 16-th January 1995.

In the following, we will concentrate on nonparametric estimates of the SPD. An extensive overview of parametric and other estimation techniques can be found, e.g., in Jackwerth (1999).

An example of option prices data set is given on the left hand side of Figure 1. You can see the typical shape of the data (decreasing and convex) for various maturities. The structure of the data can be clearly seen on the kernel regression surface which is also included in the plot. The right plot in Figure 1 displays the data only for the shortest time to expiry. In this paper, we will concentrate on this type of data and we will propose a simple method for fitting a curve satisfying all the required shape constraints such as monotonicity, convexity and the fact that the second derivative (SPD) integrates to (less than) one.

2 Construction of the estimate

In this section, we construct an estimate of the state price density satisfying all of the shape constraints which follow from the theoretical properties (no-arbitrage assumptions) of the option prices.

2.1 Notation

Let us denote the i-th observation of strike price by K_i and the corresponding option price by $C_i = C_{t,i}(K_i, T)$. In praxis, one observes option prices repeatedly for small number of distinct strike prices. Therefore, it is useful to adopt the following notation. Let $\mathcal{C} = (C_1, \ldots, C_n)^\top$ be the vector of observed option prices. We assume that the corresponding vector of the strike prices has the following structure:

$$\mathcal{K} = \begin{pmatrix} K_1 \\ K_2 \\ \vdots \\ K_n \end{pmatrix} = \begin{pmatrix} k_1 \mathbf{1}_{n_1} \\ k_2 \mathbf{1}_{n_2} \\ \vdots \\ k_p \mathbf{1}_{n_p} \end{pmatrix},$$

where $k_1 < k_2 < \cdots < k_p$, $n_j = \sum_{i=1}^{n} \mathbf{I}(K_i = k_j)$ with $\mathbf{I}(.)$ denoting the the identificator function and $\mathbf{1}_n$ vector of ones of length n.

The symbol μ_j will denote the expected value of the option price $C_t(K,T)$ in $K = k_j$.

2.2 Assumptions and constraints

We assume that the i-th observed option price (corresponding to strike price K) follows the regression model

$$C_{t,i}(K,T) = \mu(K) + \varepsilon_i, \tag{3}$$

where ε_i are i.i.d. $\mathcal{N}(0,\sigma^2)$ distributed variables. Heteroskedasticity can be incorporated in model (3) if we assume that the random errors ε_i have $\mathcal{N}(0,\sigma_K^2)$ distribution.

From the theory of option pricing it follows that the function of true conditional means $\mu(.)$ has to satisfy the following no-arbitrage constraints:

1. it is positive,
2. it is decreasing in K,
3. it is convex,
4. its second derivative exists and it is a density (i.e., nonnegative and it integrates to one).

Let us now consider the family, $\mathcal{F}$, of functions satisfying Constraints 1–4.

LEMMA 1 *Suppose that $f \in \mathcal{F}$. Then we have for its first derivative, f', that $\lim_{x\to+\infty} f'(x) = 0$ and $\lim_{x\to-\infty} f'(x) = -1$.*

PROOF:
Constraint 4 implies that the first derivative, f', exists and that it is differentiable. $\lim_{x\to+\infty} f'(x)$ exists since the function f' is increasing (Constraint 3) and bounded (Constraint 2). Next, $\lim_{x\to+\infty} f'(x) = 0$ since negative limit would violate Constraint 1 for large x ($f'(x)$ cannot be positive because $f(x)$ is decreasing). Finally, Constraint 4, $\int_{-\infty}^{\infty} f''(x)dx = 1 = \lim_{x\to+\infty} f'(x) - \lim_{x\to-\infty} f'(x)$, leads that $\lim_{x\to-\infty} f'(x) = -1$. $\square$

2.3 Existence and uniqueness

In this subsection we address the issuess of existence and uniqueness of a regression function satisfying the above stated assumptions and constraints.

In praxis, we don't deal with continuous function. Hence, we restate Constraints 1–4 for discrete functions, defined only on finite set of points, say $x_1, \ldots, x_n$, in terms of their function values, $f(x_i)$, and their scaled first differences, $f^{(1)}_{x_i,x_j} = \frac{f(x_i)-f(x_j)}{x_i-x_j}$.

5. $f(x_i) > 0$, $i = 1, \dots, n$,
6. $x_i < x_j$ implies that $f(x_i) \geq f(x_j)$,
7. $x_i < x_j < x_k$ implies that $-1 \leq f^{(1)}_{x_i,x_j} \leq f^{(1)}_{x_j,x_k} \leq 0$.

It is easy to see that Constraints 5–7 are discrete versions of Constraints 1–4.

From now on, similarly as in Robertson, Wright and Dykstra (1988), we think of the collection, $\mathcal{F}$, of functions satisfying Constraints 5–7 as a subset of a p-dimensional Euclidean space, where p is the number of distinct x_is. The constrained regression, $\hat{g}$, is in this setting the closest point of $\mathcal{F}$ to the observed g with distances measured by the usual Euclidean distance

$$d(f,g) = (f-g)^\top (f-g) = \sum_{i=1}^{n} \{f(x_i) - g(x_i)\}^2. \tag{4}$$

In this point of view, the regression function, $\hat{g}$, consists only of the values of the function in the points $x_1, \dots, x_n$. The first and second differences can be then used to approximate the first and the second derivatives, respectively.

We claim that the set, $\mathcal{F}$, of functions satisfying Constraints 5–7, has the following properties

1. $\mathcal{F}$ is closed in the topology induced by the metric given by (4),
2. $\mathcal{F}$ is convex, i.e., if $f, g \in \mathcal{F}$ and $0 \leq a \leq 1$, then $af + (1-a)g \in \mathcal{F}$.

LEMMA 2 *Assume that $\hat{g}$ is the regression of $g(x_i)$ on $x_1 \leq \dots \leq x_n$ under Constraints 5–7. If a and b are constants such that $a \leq g(x_i) \leq b$, $\forall i$, then $a - (x_n - x_1) \leq \hat{g}(x_i) \leq b + (x_n - x_1)$.*

PROOF:
It is not possible that $\hat{g}(x_i)$ lies bellow a or above b for all x_i's (otherwise we would get better fit only by shifting $\hat{g}(x_i)$). The bounds now follow from Constraint 7. $\square$

THEOREM 1 *A regression, $\hat{g} = \arg\min_{f \in \mathcal{F}} d(g,f)$, satisfying Constraints 5–7 exists.*

PROOF:
Lemma 2 implies that $\hat{g}$ belongs to a subset, $\mathcal{S}$, of $\mathcal{F}$ bounded below by $a - (x_n - x_1)$ and above by $b + (x_n - x_1)$. Thinking of the functions as of points in Euclidean space, it is clear that the continuous function $d(f,g)$ attains its minimum on the closed and bounded set $\mathcal{S}$. $\square$

REMARK 1 *Suppose $\mathcal{F}$ is any convex set of functions on $\mathcal{X}$ and g is a given function on $\mathcal{X}$. If $\hat{g} = \arg\min_{f \in \mathcal{F}} d(g,f)$ then for every $f \in \mathcal{F}$,*

$$\sum_{i=1}^{n} \{g(x_i) - \hat{g}(x_i)\}^\top \{\hat{g}(x_i) - f(x_i)\} \geq 0. \tag{5}$$

There exists at most one function $\hat{g}$ satisfying (5).

PROOF:
See Robertson, Wright and Dykstra (1988, Theorem 1.3.1). $\square$

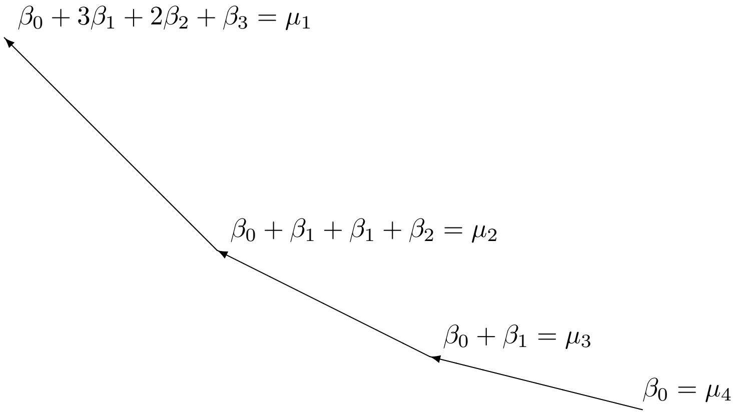

Figure 2: Illustration of the dummy variables.

COROLLARY 1 *A regression, $\hat{g}$, satisfying Constraints 5–7 exists and it is unique.*

PROOF:
It follows from Theorem 1 and Remark 1. $\square$

2.4 Regression model

The configuration of data, under Constraints 5–7 of Subsection 2.2, can be easily described using simple regression model with constraints.

In the following, we fix the time t and the expiry date T and we omit these symbols from the notation. Let us assume that the option prices $C_i(K)$ are repeatedly observed for small number p of distinct strike prices K (such setup can be seen, for example, in the right hand plot in Figure 1), where we have many observations ($n = 575$) observed only for $p = 8$ distinct strike prices.

For simplicity of the following presentation we display the coefficients β_i in the situation with only four distinct strike prices ($p = 4$) in Figure 2.

Defining the expected values of the option prices given strike price, $\mu_j = EC(K_j)$, we can write

$$\begin{aligned}
\mu_p &= \beta_0, \\
\mu_{p-1} &= \beta_0 + \beta_1, \\
\mu_{p-2} &= \beta_0 + 2\beta_1 + \beta_2, \\
\mu_{p-3} &= \beta_0 + 3\beta_1 + 2\beta_2 + \beta_3, \\
&\vdots \\
\mu_1 &= \beta_0 + (p-1)\beta_1 + (p-2)\beta_2 + \cdots + \beta_{p-1}.
\end{aligned}$$

Thus, we fit our data using coefficients β_j, $j = 1, \ldots, p$. The conditional means μ_i, $i = 1, \ldots, p$ are replaced by the same number of parameters β_j, $j = 0, \ldots, p-1$ which allow to impose the shape constraints in a more natural way.

The interpretation of the coefficients β_j can be seen from Figure 2. β_0 is the mean option price at point 4. Constraint 1, Subsection 2.2, implies that it has to be positive. β_1 is the difference between the mean option prices at point 4 and point 3; Constraint 2 implies that it has to be positive. The next coefficient, β_2, approximates change in first derivative in point 3 and it can be interpreted as an approximation of the second derivative in point 3. Constraint 3 implies that β_2 has to be positive. Similarly, β_3 is an estimate of the (positive) second derivative of $C(K)$ in point 2. Constraint 4 can be rewritten as $\beta_2 + \beta_3 \leq 1$.

In praxis, we start with the construction of a design matrix which allows us to write the above model in the following linear form. For simplicity of presentation, we again set $p = 4$:

$$\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 2 & 1 \\ 1 & 2 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}. \tag{6}$$

Ignoring the constraints on the coefficients would lead to simple linear regression problem. Unfortunatelly, this approach does not have to lead, and usually does not, to reasonable results.

Model (6) in the above form can be reasonably interpreted only if the observed strike prices are equidistant and if the distances between the neigbouring observed strike prices are equal to one. If we want to keep the interpretation of the parameters β_j as the derivatives of the estimated function, we should use the design matrix

$$\Delta = \begin{pmatrix} 1 & \Delta_p^1 & \Delta_{p-1}^1 & \Delta_{p-2}^1 & \cdots & \Delta_3^1 & \Delta_2^1 \\ 1 & \Delta_p^2 & \Delta_{p-1}^2 & \Delta_{p-2}^2 & \cdots & \Delta_3^2 & 0 \\ \vdots & & & & & & \vdots \\ 1 & \Delta_p^{p-2} & \Delta_{p-1}^{p-2} & 0 & \cdots & 0 & 0 \\ 1 & \Delta_p^{p-1} & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \tag{7}$$

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

The vector of conditional means μ can be written in terms of the parameters β as follows

$$\begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{pmatrix} = \mu = \Delta\beta = \Delta \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{p-1} \end{pmatrix}. \tag{8}$$

The constraints on the conditional means μ_j can now be expressed as conditions on the parameters of the model (8). Namely, it suffices to request that $\beta_i > 0$, $i = 0, \ldots, p-1$ and that $\sum_{j=2}^{p-1} \beta_j \leq 1$.

The model for the observed option prices $\mathcal{C}$ and for the observed strike prices K becomes

$$\mathcal{C} = \mathcal{X}_\Delta \beta + \varepsilon, \tag{9}$$

where $\mathcal{X}_\Delta$ is the design matrix in which each row of the matrix Δ is repeated n_j times, $j = 1, \ldots, p$.

2.5 Implementing the constraints

In order to impose Constraints 5–7 on parameters β_i, $i = 0, \dots, p-1$, we propose the following parametrization of the model in terms of parameters θ_j, $j = 0, \dots, p$:

$$\begin{aligned}
\beta_0(\theta) &= \exp(\theta_0), \\
\beta_1(\theta) &= \frac{\exp(\theta_1)}{\sum_{j=1}^{p} \exp(\theta_j)}, \\
&\vdots \\
\beta_{p-1}(\theta) &= \frac{\exp(\theta_{p-1})}{\sum_{j=1}^{p} \exp(\theta_j)}.
\end{aligned}$$

Clearly, parameters $\beta_i(\theta)$ satisfy the constraints

$$\begin{aligned}
\beta_i(\theta) &> 0, \quad i = 0, \dots, p-1, \\
\sum_{j=2}^{p-1} \beta_j(\theta) &< 1.
\end{aligned}$$

This means that the parameters $\beta_2(\theta), \dots, \beta_{p-1}(\theta)$ can be considered as point estimates of the state price density (the estimates have to be positive and integrate to less than one). Furthermore, in view of Lemma 1, it is worthwhile to note that the parameters satisfy also

$$-\sum_{j=1}^{k} \beta_j \in (-1, 0), \quad \text{for } k = 1, \dots, p-1.$$

Notice that $\sum_{j=1}^{k} \beta_j$, for $k = p-1, \dots, 1$, can be interpreted as estimates of the integrated state price density.

The equality

$$\exp(\theta_p) \left\{ \sum_{j=1}^{p-1} \exp(\theta_j) \right\}^{-1} = 1 - \left\{ \sum_{j=1}^{p-1} \beta_j(\theta) \right\}^{-1}$$

shows the meaning of the parameter θ_p. Setting this parameter to $-\infty$ is the same as requiring that $\sum_{j=2}^{p-1} \beta_j(\theta) = 1$. This also allows to test the hypothesis whether our data cover the support of the state price density as H_0: $\theta_p = \infty$ against H_1: $\theta_p \neq -\infty$.

The model (9) written in terms of parameters θ_i, $i = 0, \dots, p$ is a nonlinear regression model which can be estimated using standard maximum likelihood methods. The main advantage of the maximum likelihood estimator (MLE) is that the asymptotic distribution is well known and that the asymptotic variance of the estimator can be approximated using numerical methods implemented in many statistical packages.

Using the data displayed in the right hand plot in Figure 1, we obtain the estimates displayed in Figure 3. The top plot displays the original data, the second plot shows the estimate of the first derivative, and the third plot shows the estimate of the second derivative, i.e., the state price density. Actually, all plots contain two curves, both obtained using model (9). The thick line is calculated using the parameters β_i without constraints whereas the thin line uses the reparametrization $\beta_i(\theta)$ given in Subsection 2.5. In Figure 3, these two estimates coincide since the

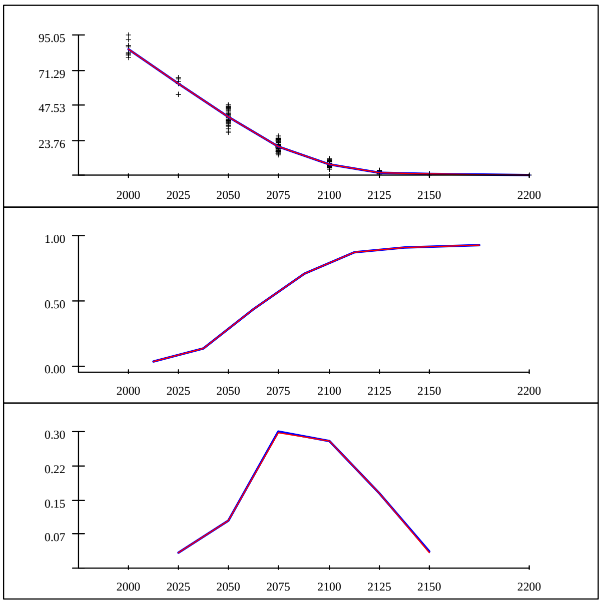

Figure 3: 16th January 1995

model maximizing the likelihood without constraints, by chance, fulfills the constraints ($\exists \theta : \beta_i = \beta_i(\theta)$, $i = 0, \dots, p-1$) and hence it is clear that the same parameters maximize also the constrained likelihood.

The situation, in which the estimates with and without constraints differ, is displayed in Figure 4. Notice that the difference between the two regression curves is small whereas the difference between the estimates of the state price density (i.e., the second derivative of the curve) is surprisingly large. The unconstrained estimate shows very bad behaviour on the left hand side of the plot. The constrained version behaves much more reasonably. Very small difference between the fitted lines in the top plot in Figure 4 leads to huge differences in the estimates of second derivative.

We can conclude that small errors in the estimates of the curve can lead to huge errors in the estimates of the first and second derivatives. The scale of this type of error seems to be limited by imposing the shape constraints given in Subsection 2.2.

2.6 Inverse transformation of model parameters

For the numerical algorithm, it is useful to know how to calculate θs from given βs. This is needed, for example, to obtain reasonable starting point for the iterative procedure maximizing the likelihood.

LEMMA 3 *Given $\beta = (\beta_1, \dots, \beta_p)^\top$, where $\beta_p = 1 - \sum_{i=1}^{p-1} \beta_i$, the parameters $\theta = (\theta_1, \dots, \theta_p)^\top$ satisfy the system of equations*

$$\left(\beta \mathbf{1}_p^\top - \mathbf{I}_p\right) \exp \theta^\top = \mathcal{A} \exp \theta^\top = 0, \tag{10}$$

where $\mathbf{1}$ denotes vector of ones and $\mathbf{I}$ is the identity matrix. Furthermore,

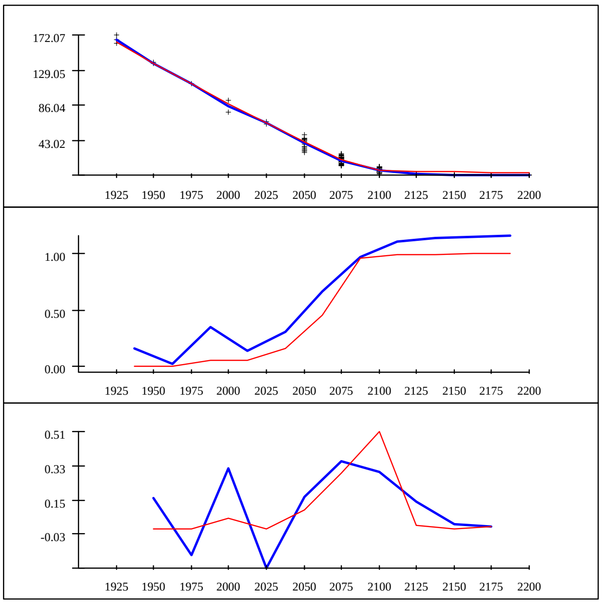

Figure 4: 17th January 1995

$$\operatorname{rank} \mathcal{A} = p - 1. \tag{11}$$

The system of equations (10) has infinitely many solutions which can be expressed as

$$\exp(\theta) = \left(\mathcal{A}^{-}\mathcal{A} - \mathbf{I}_p\right) z, \tag{12}$$

where $\mathcal{A}^{-}$ denotes the generalized inverse of $\mathcal{A}$ and where z is an arbitrary vector in $\mathbb{R}^p$ such that the right hand side of (12) is positive.

PROOF:
Parts (10) and (11) follow from the definition of $\beta(\theta)$ and from simple algebra (notice that the sum of rows of $\mathcal{A}$ is equal to zero). Part (12) follows, e.g., from Anděl (1985, Theorem IV.18). $\square$

It remains to choose the vector z in (12) so that the solution of the system of equations (10) is positive.

PROPOSITION 1 *The rank of matrix $\mathcal{A}^{-}\mathcal{A} - \mathbf{I}_p$ is 1. Hence, any solution of the system of equations (10) is a multiple of the first column of the matrix $\mathcal{A}^{-}\mathcal{A} - \mathbf{I}_p$. The vector z in (12) can be chosen, e.g., as $z = \pm\mathbf{1}_p$, where the sign is chosen so that the resulting solution is positive.*

PROOF:
The definition of the generalized inverse is

$$\mathcal{A}\mathcal{A}^{-}\mathcal{A} - \mathcal{A} = \mathcal{A}(\mathcal{A}^{-}\mathcal{A} - \mathbf{I}_p) = 0. \tag{13}$$

Lemma 3 says that $\operatorname{rank} \mathcal{A} = p - 1$. Hence, equation (13) implies that $\operatorname{rank}(\mathcal{A}^{-}\mathcal{A} - \mathbf{I}_p) \leq 1$. Noticing that $\mathcal{A}^{-}\mathcal{A} \neq \mathbf{I}_p$ means that $\operatorname{rank}(\mathcal{A}^{-}\mathcal{A} - \mathbf{I}_p) > 0$ and concludes the proof. $\square$

2.7 The algorithm

The proposed algorithm consists of the following steps:

1. obtain reasonable initial estimate $\hat{\beta}$, e.g., by running PAV algorithm on the unconstrained least squares estimates of the first derivative of the curve,
2. transform the initial estimates $\hat{\beta}$ into the estimates $\hat{\theta}$ using the method described in Subsection 2.6,
3. minimize the nonlinear least squares as described in Subsection 2.5 using numerical methods.

An application of this simple algorithm on real data is given in the next section.

3 Application on DAX data

In order to illustrate the method, we apply it on observed DAX option prices on two consecutive days. These days (16th and 17th January 1995) were selected since they provide nice insight into the behaviour of the presented methods.

The observed option prices on one day (16th January 1995) are plotted on the left hand side of Figure 1 against maturity and strike price. The shape of dependency of the option price on the strike price can be nicely observed. For simplicity, in the following analyses we restrict ourselves only to data for fixed maturity as displayed on the right plot in Figure 1.

In Figures 3 and 4 we observe the difference between the unconstrained linear regression estimate and the constrained nonlinear regression estimate described in Section 2.5.

In Figure 3, the unconstrained model incindentally satisfies all conditions on the shape of the curve. Hence, the estimates of the curve itself (1st plot) its first derivative (2nd plot) and the SPD (3rd plot) coincide.

On 17th January, the situation becomes more interesting and it illustrates very clearly the advantages of the constrained estimator. In Figure 4, we plot the unconstrained and the constrained estimates using thick and thin line, respectively. Clearly, the difference between the fitted data (1st plot) is very small. However, this small difference in the first plot results in huge differences in the estimate of the first derivative (2nd plot) and especially in the estimate of the second derivate, the SPD, in the 3rd plot of Figure 4.

Figures 3 and 4 are, in more detail, discussed also in Subsection 2.5.

Interpretation of the estimates

The coefficients, $\hat{\beta}_{p-1}, \dots, \hat{\beta}_2$, plotted in Figures 3 and 4 can be described as estimates of the changes of the first derivative in that point. Since the first derivative of the curve corresponds to the integrated SPD, the coefficients $\hat{\beta}_{p-1}, \dots, \hat{\beta}_2$ estimate probabilities associated with the corresponding strike price. These estimates can thus be interpreted as histogram-like estimates of the state price density. Obviously, the next step might be kernel smoothing of the above estimate which would easily provide continuous and smooth estimates of the SPD. Using asymptotic distribution of the Maximum Likelihood estimates we can obtain asymptotic pointwise confidence intervals for the smoothed curve.

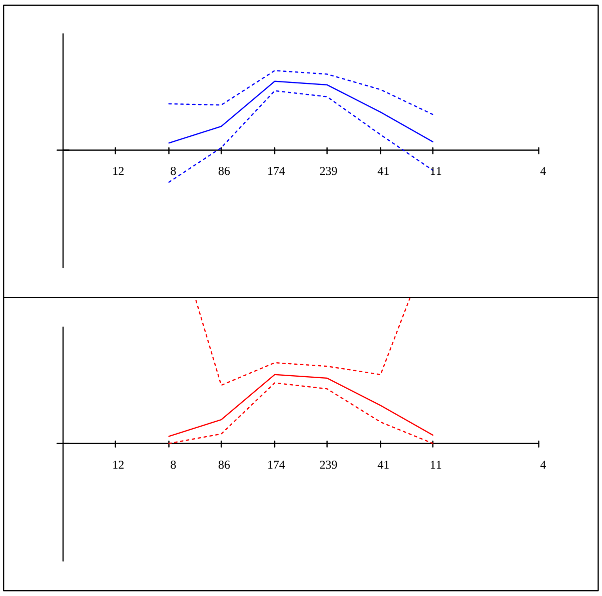

Figure 5: Confidence intervals for SPD on 16th January 1995.

Using this model, it might be interesting to test whether the parametric models commonly used in praxis (e.g., mixtures of log-normal distributions) are consistent with our nonparametric estimate of the SPD.

Confidence intervals for SPD

We present two simple methods for calculating pointwise confidence intervals for the SPD. The description of the x-axis in Figures 5 and 6 shows the number of observations at each of the design points.

Notice that, in the unconstrained model, the estimates of the values of the SPD are directly the parameters of the linear regression model. Hence, the confidence intervals for the parameters are also confidence intervals for the SPD. These confidence intervals for 16th and 17th January are displayed in upper plots in Figures 5 and 6. The drawbacks of this method are clearly visible. In Figure 5, the lower bounds of the confidence intervals do not satisfy the condition of positivity. In Figure 6, we observe large variability on the left-hand side of the plot (the region with low number of observations). Again, some of the lower bounds are not positive.

Clearly, the confidence intervals based on the unconstrained model makes sense only if the constraints are, by chance, satisfied. Even if this is the case, there is no guarantee that the lower bounds will be positive.

The lower plots in Figures 5 and 6 display confidence intervals conditional on the fact that $\sum_{i=1}^{p} \exp(\theta_i) = 1$. This conditioning corresponds

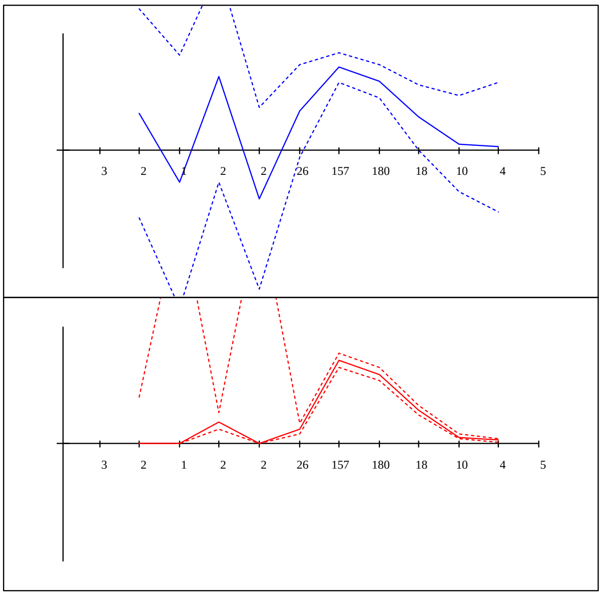

Figure 6: Confidence intervals for SPD on 17th January 1995.

to the following parametrization of the model:

$$
\begin{aligned}
\beta_0(\theta) &= \exp(\theta_0), \\
\beta_1(\theta) &= \exp(\theta_1), \\
\beta_2(\theta) &= \exp(\theta_2), \\
&\vdots \\
\beta_{p-1}(\theta) &= \exp(\theta_{p-1}),
\end{aligned}
$$

under the constraint that $\sum_{i=1}^{p-1} \beta_i(\theta) < 1$. Using maximum likelihood theory, we calculate confidence intervals for the parameters θ (rescaled so that $\sum_{i=1}^{p} \exp(\theta_i) = 1$). Exponentiating the limits of these confidence intervals leads to valid confidence intervals for parameters β.

In Figure 5, both type of confidence intervals provide very similar results. The only difference is at the minimum and maximum value of the independent variable (strike price) where the unconstrained method provides negative lower bounds and the conditional method leads to very large upper bounds of the confidence intervals.

In Figure 6, we plot the confidence intervals for January 17th. Here, the unconstrained and the conditional methods lead to very different estimates. We can observe that the confidence intervals on the right hand side are much narrower for the conditional method. On the left hand side, both methods tend to provide confidence intervals that look very wide. For the conditional method, we observe that the confidence intervals look "suspicious" when the estimated value of the SPD is very close too zero and when the number of observation in that region (see the description of the x-axis) is small.

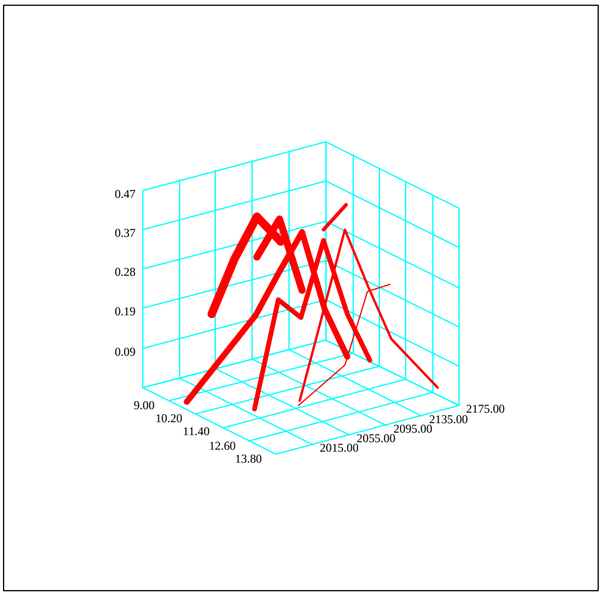

Figure 7: Estimate of SPD on 17th January 1995 for each hour.

Further considerations

The SPD estimates in Figure 7 were calculated for each hour separately in order to display the development of the SPD during the day. The thickness of the line corresponds to the time coordinate. Clearly, the estimate does not change "too much" during the day. It might be easily tested if there is a change in the behaviour of the SPD at a given time. Methods for identifying a change of behaviour (changepoint) are one of the topics for further research.

Appendix

Maximum Likelihood

Assuming normality, the log-likelihood for the model (9) can be written as

$$l(\mathcal{C}, \mathcal{X}_\Delta, \theta, \sigma) = -n \log \sigma - \frac{1}{2\sigma^2}\{\mathcal{C} - \mathcal{X}_\Delta \beta(\theta)\}^\top \{\mathcal{C} - \mathcal{X}_\Delta \beta(\theta)\}, \tag{14}$$

where $\mathcal{X}_\Delta$ is the design matrix given in (9). The maximum likelihood estimator is defined as

$$\hat{\theta} = \arg\max_\theta l(\mathcal{C}, \mathcal{X}_\Delta, \theta, \sigma). \tag{15}$$

In order to implement the described algorithm numerically, it is useful to express the contribution of the i-th row to the log-likelihood in the following form:

$$
\begin{aligned}
l_i(\theta) &= -\log\sigma - \frac{1}{2\sigma^2}\Bigg\{C_i(K) - \exp(\theta_0) - \Delta_p^i \frac{\exp(\theta_1)}{\sum_{j=1}^p \exp(\theta_j)} \\
&\quad - \Delta_{p-1}^i \frac{\exp(\theta_2)}{\sum_{j=1}^p \exp(\theta_j)} - \cdots - \Delta_2^i \frac{\exp(\theta_{p-1})}{\sum_{j=1}^p \exp(\theta_j)}\Bigg\}^2 \\
&= -\log\sigma - \frac{1}{2\sigma^2} r_i^2, \qquad (16)
\end{aligned}
$$

where $\Delta_i^j = 0$ if $j \leq i$ and where r_i denotes the i-th residual. The derivative of (16) with respect to the unknown parameters $\theta_0, \ldots, \theta_p$ is

$$
\begin{aligned}
\frac{\partial l_i(\theta)}{\partial \theta_0} &= \frac{1}{\sigma^2} r_i \exp(\theta_0), \\
\frac{\partial l_i(\theta)}{\partial \theta_k} &= -\frac{1}{\sigma^2} r_i \Bigg\{ -\Delta_p^i \frac{\exp(\theta_1)\exp(\theta_k)}{\left\{\sum_{j=1}^p \exp(\theta_j)\right\}^2} \\
&\quad - \cdots + \Delta_{p-k+1}^i \frac{\exp(\theta_k)\left\{\sum_{j=1}^p \exp(\theta_j) - \exp(\theta_k)\right\}}{\left\{\sum_{j=1}^p \exp(\theta_j)\right\}^2} \\
&\quad - \cdots - \Delta_2^i \frac{\exp(\theta_{p-1})\exp(\theta_k)}{\left\{\sum_{j=1}^p \exp(\theta_j)\right\}^2} \Bigg\} \\
&= -\frac{1}{\sigma^2} r_i \beta_k(\theta) \left\{ \Delta_{p-k+1}^i - (\Delta_p^i, \ldots, \Delta_2^i) \begin{pmatrix} \beta_1(\theta) \\ \vdots \\ \beta_{p-1}(\theta) \end{pmatrix} \right\}.
\end{aligned}
$$

References

Aït-Sahalia, Y. & Duarte, J. (2003). Nonparametric Option Pricing under Shape Restrictions, *Journal of Econometrics*, **116**, 9–47.

Aït-Sahalia, Y. & Lo, A.W. (1998). Nonparametric estimation of state-price densities implicit in financial asset prices, *Journal of Finance* **53**, 499–547.

Aït-Sahalia, Y., Wang, Y. & Yared, F. (2000). Do Option Markets Correctly Price the Probabilities of Movement of the Underlying Asset?, *Journal of Econometrics* **102**, 67–110.

Anděl, J. (1985). *Mathematical Statistics (in Czech)*, SNTL/Alfa, Prague.

Breeden, D. & Litzenberger, R. (1978). Prices of state-contingent claims implicit in option prices, *Journal of Business* **51**, 621–651.

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

Härdle, W. & Yatchew, A. (2001). Nonparametric state price density estimation using constrained least squares and the bootstrap, *Discussion paper*, Department of Economics, University of Toronto.

Robertson, T., Wright, F.T. & Dykstra, R.L. (1988). *Order Restricted Statistical Inference*, Wiley, Chichester.