Detlefsen, Kai; Härdle, Wolfgang Karl

Working Paper

Common functional implied volatility analysis

SFB 649 Discussion Paper, No. 2005-012

Provided in Cooperation with:
Collaborative Research Center 649: Economic Risk, Humboldt University Berlin

Suggested Citation: Detlefsen, Kai; Härdle, Wolfgang Karl (2005) : Common functional implied volatility analysis, SFB 649 Discussion Paper, No. 2005-012, Humboldt University of Berlin, Collaborative Research Center 649 - Economic Risk, Berlin

This Version is available at:
https://hdl.handle.net/10419/25031

SFB 649 Discussion Paper 2005-012

Common Functional Implied Volatility Analysis

Michal Benko*
Wolfgang Härdle*

* CASE - Center for Applied Statistics and Economics,
Humboldt-Universität zu Berlin, Germany

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

http://sfb649.wiwi.hu-berlin.de
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin

1 Common functional implied volatility analysis

Michal Benko and Wolfgang Härdle

ACKNOWLEDGEMENT

Financial support of *Deutsche Forschungsgemeinschaft* via SFB 649 "Ökonomisches Risiko", Humboldt-Universität zu Berlin, is gratefully acknowledged.

JEL classification: C13, G19

1.1 Introduction

Trading, hedging and risk analysis of complex option portfolios depend on accurate pricing models. The modelling of implied volatilities (IV) plays an important role, since volatility is the crucial parameter in the Black-Scholes (BS) pricing formula. It is well known from empirical studies that the volatilities implied by observed market prices exhibit patterns known as volatility smiles or smirks that contradict the assumption of constant volatility in the BS pricing model. On the other hand, the IV is a function of two parameters: the strike price and the time to maturity and it is desirable in practice to reduce the dimension of this object and characterize the IV surface through a small number of factors. Clearly, a dimension reduced pricing-model that should reflect the dynamics of the IV surface needs to contain factors and factor loadings that characterize the IV surface itself and their movements across time.

A popular dimension reduction technique is the principal components analysis (PCA), employed for example by Fengler, Härdle, and Schmidt (2002) in the IV surface analysis. The discretization of the strike dimension and application of PCA yield suitable factors (weight vectors) in the multivariate framework. Noting that the IVs of fixed maturity could also be viewed as random functions,

we propose to use the functional analogue of PCA. We utilize the truncated functional basis expansion described in Ramsay and Silverman (1997) to the IVs of the European options on the German stock index (DAX). The standard functional PCA, however, yields weight functions that are too rough, hence a smoothed version of functional PCA is proposed here.

Like Fengler, Härdle, and Villa (2003) we discover similarities of the resulting weight functions across maturity groups. Thus we propose an estimation procedure based on the Flury-Gautschi algorithm, Flury (1988), for the simultaneous estimation of the weight functions for two different maturities. This procedure yields common weight functions with the level, slope, and curvature interpretation known from the financial literature. The resulting common factors of the IV surface are the basic elements to be used in applications, such as simulation based pricing, and deliver a substantial dimension reduction.

The chapter is organized as follows. In Section 1.2 the basic financial framework is presented, while in Section 1.3 we introduce the notation of the functional data analysis. In the following three sections we analyze the IV functions using functional principal components, smoothed functional principal components and common estimation of principal components, respectively.

1.2 Implied volatility surface

Implied volatilities are derived from the BS pricing formula for European options. Recall that European call and put options are derivatives written on an underlying asset S driven by the price process S_t, which yield the pay-off $\max(S_T - K, 0)$ and $\max(K - S_T, 0)$ respectively, at a given expiry time T and for a prespecified strike price K. The difference $\tau = T - t$ between the day of trade and day of expiration (maturity) is called time to maturity. The pricing formula for call options, Black and Scholes (1973), is:

$$\begin{aligned} C_t(S_t, K, \tau, r, \sigma) &= S_t \Phi(d_1) - K e^{-r\tau} \Phi(d_2) \\ d_1 &= \frac{\ln(S_t/K) + (r + 1/2\sigma^2)\tau}{\sigma\sqrt{\tau}},\ d_2 = d_1 - \sigma\sqrt{\tau}, \end{aligned} \tag{1.1}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution, r is the riskless interest rate, and σ is the (unknown and constant) volatility parameter. The put option price P_t can be obtained from the put-call parity $P_t = C_t - S_t + e^{-\tau r} K$.

For a European option the implied volatility $\hat{\sigma}$ is defined as the volatility – σ, which yields the BS price C_t equal to the price $\tilde{C}_t$ observed on the market. For a single asset, we obtain at each time point t a two-dimensional function – the IV surface $\hat{\sigma}_t(K, \tau)$. In order to standardize the volatility functions in time, one may rescale the strike dimension by dividing K by the future price $F_t(\tau)$ of the underlying asset with the same maturity. This yields the so-called moneyness $\kappa = K/F_t(\tau)$. Note that some authors define moneyness simply as $\kappa = K/S_t$. In contrast to the BS assumptions, empirical studies show that IV surfaces are significantly curved, especially across the strikes. This phenomenon is called a volatility smirk or smile. Smiles stand for U-shaped volatility functions and smirks for decreasing volatility functions.

We focus on the European options on the German stock index (ODAX). Figure 1.1 displays the ODAX implied volatilities computed from the BS formula (red points) and the IV surface on May 24, 2001 estimated using a local polynomial estimator for $\tau \in [0, 0.6]$ and $\kappa \in [0.8, 1.2]$. We can clearly observe the "strings" of the original data on maturity grid $\tau \in \{0.06111, 0.23611, 0.33333, 0.58611\}$, which corresponds to 22, 85, 120, and 211 days to maturity. This maturity grid is structured by market conventions and changes over time. The fact that the number of transactions with short maturity is much higher than those with longer maturity is also typical for the IVs observed on the market.

The IV surface is a high-dimensional object – for every time point t we have to analyze a two-dimensional function. Our goal is to reduce the dimension of this problem and to characterize the IV surface through a small number of factors. These factors can be used in practice for risk management, e.g. with vega-strategies.

The analyzed data, taken from MD*Base, contains EUREX intra-day transaction data for DAX options and DAX futures (FDAX) from January 2 to June 29, 2001. The IVs are calculated by the Newton-Raphson iterative method. The correction of Hafner and Wallmeier (2001) is applied to avoid influence of the tax-scheme in the DAX. For robustness, we exclude the contracts with time to maturity of less than 7 days and maturity strings with less than 100 observations. The approximation of the "riskless" interest rate with a given maturity is obtained on a daily basis from the linear interpolation of the 1, 3, 6, and 12 month EURIBOR interest rates (obtained from Datastream).

The resulting data set is analyzed using the functional data analysis framework. One advantage of this approach, as we will see later in this chapter, is the possibility of introducing smoothness in the functional sense and using it for regularization. The notation of the functional data analysis is rather complex,

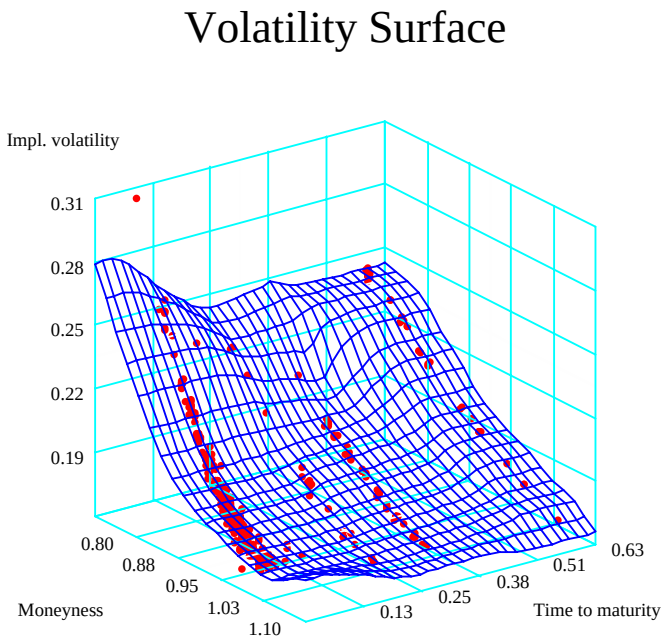

Figure 1.1: Implied volatility surface of ODAX on May 24, 2001.

STFfda01.xpl

therefore the theoretical motivation and the basic notation will be introduced in the next section.

1.3 Functional data analysis

In the functional data framework, the objects are usually modelled as realizations of a stochastic process $X(t), t \in J$, where J is a bounded interval in $\mathbb{R}$. Thus, the set of functions

$$x_i(t),\ i = 1, 2, \ldots n,\ t \in J,$$

represents the data set. We assume the existence of the mean, variance, and covariance functions of the process $X(t)$ and denote these by $\mathrm{E}X(t)$, $\mathrm{Var}(t)$ and $\mathrm{Cov}(s,t)$ respectively.

For the functional sample we can define the sample-counterparts of $\mathrm{E}X(t)$, $\mathrm{Var}(t)$ and $\mathrm{Cov}(s,t)$ in a straightforward way:

$$\begin{aligned}
\bar{X}(t) &= \tfrac{1}{n}\sum_{i=1}^{n} x_i(t), \\
\widehat{\mathrm{Var}}(t) &= \tfrac{1}{n-1}\sum_{i=1}^{n} \left\{x_i(t) - \bar{X}(t)\right\}^2, \\
\widehat{\mathrm{Cov}}(s,t) &= \tfrac{1}{n-1}\sum_{i=1}^{n} \left\{x_i(s) - \bar{X}(s)\right\}\left\{x_i(t) - \bar{X}(t)\right\}.
\end{aligned}$$

In practice, we observe the function values $\mathcal{X} \stackrel{\mathrm{def}}{=} \{x_i(t_{i1}), x_i(t_{i2}), \ldots, x_i(t_{ip_i}); i = 1, \ldots, n\}$ only on a discrete grid $\{t_{i1}, t_{i2}, \ldots, t_{ip_i}\} \in J$, where p_i is the number of grid points for the ith observation. One may estimate the functions $x_1, \ldots, x_n$ via standard nonparametric regression methods, Härdle (1990). Another popular way is to use a truncated functional basis expansion. More precisely, let us denote a functional basis on the interval J by $\{\Theta_1, \Theta_2, \ldots, \}$ and assume that the functions x_i are approximated by the first L basis functions Θ_l, $l = 1, 2, \ldots, L$:

$$x_i(t) = \sum_{l=1}^{L} c_{il}\Theta_l(t) = \mathbf{c}_i^\top \boldsymbol{\Theta}(t), \tag{1.2}$$

where $\boldsymbol{\Theta} = (\Theta_1, \ldots, \Theta_L)^\top$ and $\mathbf{c}_i = (c_{i1}, \ldots, c_{iL})^\top$. The number of basis functions L determines the tradeoff between data fidelity and smoothness. The analysis of the functional objects will be implemented through the coefficient matrix

$$\mathbf{C} = \{c_{il},\ i = 1, \ldots, n,\ l = 1, \ldots, L\}.$$

The mean, variance, and covariance functions are calculated by:

$$\begin{aligned}
\bar{X}(t) &= \bar{\mathbf{c}}^\top \boldsymbol{\Theta}(t), \\
\widehat{\mathrm{Var}}(t) &= \boldsymbol{\Theta}(t)^\top \mathrm{Cov}(\mathbf{C})\boldsymbol{\Theta}(t), \\
\widehat{\mathrm{Cov}}(s,t) &= \boldsymbol{\Theta}(s)^\top \mathrm{Cov}(\mathbf{C})\boldsymbol{\Theta}(t),
\end{aligned}$$

where $\bar{\mathbf{c}}_l \stackrel{\mathrm{def}}{=} \frac{1}{n}\sum_{i=1}^{n} c_{il},\ l = 1, \ldots, L$ and $\mathrm{Cov}(\mathbf{C}) \stackrel{\mathrm{def}}{=} \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{c}_i - \bar{\mathbf{c}})(\mathbf{c}_i - \bar{\mathbf{c}})^\top$.

The scalar product in the functional space is defined by:

$$\langle x_i, x_j \rangle \stackrel{\text{def}}{=} \int_J x_i(t)x_j(t)dt = \mathbf{c}_i^\top \mathbf{W} \mathbf{c}_j,$$

where

$$\mathbf{W} \stackrel{\text{def}}{=} \int_J \mathbf{\Theta}(t)\mathbf{\Theta}(t)^\top dt. \tag{1.3}$$

In practice, the coefficient matrix $\mathbf{C}$ needs to be estimated from the data set $\mathcal{X}$.

An example for a functional basis is the Fourier basis defined on J by:

$$\Theta_l(t) = \begin{cases} 1, & l = 0, \\ \sin(r\omega t), & l = 2r-1, \\ \cos(r\omega t), & l = 2r, \end{cases}$$

where the frequency ω determines the period and the length of the interval $|J| = 2\pi/\omega$. The Fourier basis defined above can be easily transformed to the orthonormal basis, hence the scalar-product matrix in (1.3) is simply the identity matrix.

Our aim is to estimate the IV-functions for fixed τ = 1 month (1M) and 2 months (2M) from the daily-specific grid of the maturities. We estimate the Fourier coefficients on the moneyness-range $\kappa \in [0.9, 1.1]$ for maturities observed on particular day i. For $\tau^* = 1\text{M}, 2\text{M}$ we calculate $\hat{\sigma}_i(\kappa, \tau^*)$ by linear interpolation of the closest observable IV string with $\tau \leq \tau^*$, $\widehat{\sigma}_i(\kappa, \tau^*_{i-})$ and $\tau \geq \tau^*$, $\widehat{\sigma}_i(\kappa, \tau^*_{i+})$:

$$\hat{\sigma}_i(\kappa, \tau^*) = \hat{\sigma}_i(\kappa, \tau^*_{i-}) \left(1 - \frac{\tau^* - \tau^*_{i-}}{\tau^*_{i+} - \tau^*_{i-}}\right) + \hat{\sigma}_i(\kappa, \tau^*_{i+}) \left(\frac{\tau^* - \tau^*_{i-}}{\tau^*_{i+} - \tau^*_{i-}}\right),$$

for i where τ^*_{i-} and τ^*_{i-} exist. In Figure 1.2 we show the situation for τ^* =1M on May 30, 2001. The blue points and the blue finely dashed curve correspond to the transactions with τ^*_- =16 days and the green points and the green dashed curve to the transactions with $\tau^*_+ = 51$ days. The solid black line is the linear interpolation at $\tau^* = 30$ days.

The choice of $L = 9$ delivers a good tradeoff between flexibility and smoothness of the strings. At this moment we exclude from our analysis those days, where this procedure cannot be performed due to the complete absence of the needed maturities, and strings with poor performance of estimated coefficients, due to the small number of contracts in a particular string or presence

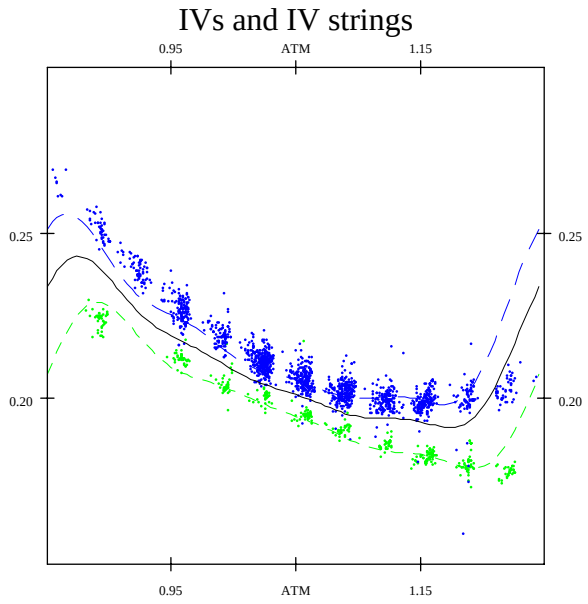

Figure 1.2: Linear interpolation of IV strings on May 30, 2001 with $L = 9$.
STFfda02.xpl

of strong outliers. Using this procedure we obtain 77 "functional" observations $x_{i_1}^{1M}(\kappa) \stackrel{\text{def}}{=} \hat{\sigma}_{i_1}(\kappa, 1M),\ i_1 = 1, \ldots, 77$, for the 1M-maturity and 66 observations $x_{i_2}^{2M}(\kappa) \stackrel{\text{def}}{=} \hat{\sigma}_{i_2}(\kappa, 2M),\ i_2 = 1, \ldots, 66$, for the 2M-maturity, as displayed in Figure 1.3.

1.4 Functional principal components

Principal Components Analysis yields dimension reduction in the multivariate framework. The idea is to find normalized weight vectors $\gamma_m \in \mathbb{R}^p$, for which the linear transformations of a p-dimensional random vector $\mathbf{x}$, with $\mathrm{E}[\mathbf{x}] = 0$:

$$f_m = \gamma_m^\top \mathbf{x} = \langle \gamma_m, \mathbf{x} \rangle,\ m = 1, \ldots, p, \tag{1.4}$$

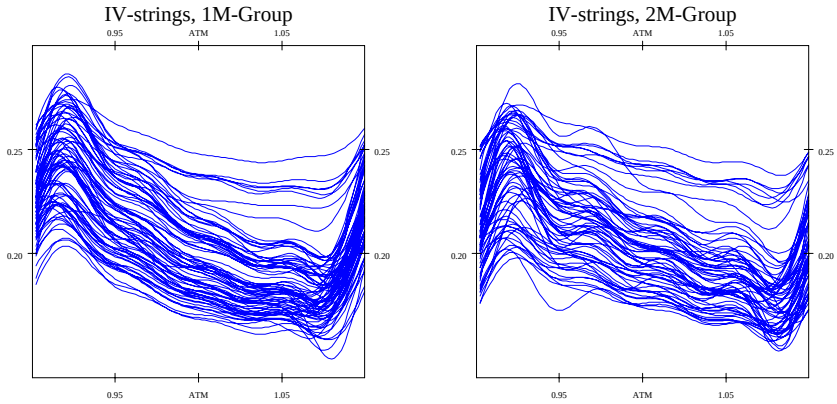

Figure 1.3: Functional observations estimated using Fourier basis with $L = 9$, $\hat{\sigma}_{i_1}(\kappa, 1M)$, $i_1 = 1, \ldots, 77$, in the left panel, $\hat{\sigma}_{i_2}(\kappa, 2M)$ $i_2 = 1, \ldots, 66$ in the right panel.

STFfda03.xpl

have maximal variance subject to:

$$\gamma_l^\top \gamma_m = \langle \gamma_l, \gamma_m \rangle = I(l = m) \text{ for } l \leq m.$$

Where I denotes the identificator function. The solution is the Jordan spectral decomposition of the covariance matrix, Härdle and Simar (2003).

In the Functional Principal Components Analysis (FPCA) the dimension reduction can be achieved via the same route, i.e. by finding orthonormal weight functions $\gamma_1, \gamma_2, \ldots$, such that the variance of the linear transformation is maximal. In order to keep notation simple we assume $\mathrm{E}X(t) = 0$. The weight functions satisfy:

$$\begin{aligned} ||\gamma_m||^2 &= \int \gamma_m(t)^2 dt = 1, \\ \langle \gamma_l, \gamma_m \rangle &= \int \gamma_l(t)\gamma_m(t)dt = 0, \ l \neq m. \end{aligned}$$

The linear transformation is:

$$f_m = \langle \gamma_m, X \rangle = \int \gamma_m(t) X(t) dt,$$

and the desired weight functions solve:

$$\underset{\langle \gamma_l, \gamma_m \rangle = I(l=m), l \leq m}{\arg\max} \operatorname{Var} \langle \gamma_m, X \rangle, \tag{1.5}$$

or equivalently:

$$\underset{\langle \gamma_l, \gamma_m \rangle = I(l=m), l \leq m}{\arg\max} \int \int \gamma_m(s) \operatorname{Cov}(s,t) \gamma_m(t) ds dt.$$

The solution is obtained by solving the Fredholm functional eigenequation

$$\int \operatorname{Cov}(s,t) \gamma(t) dt = \lambda \gamma(s). \tag{1.6}$$

The eigenfunctions $\gamma_1, \gamma_2, \ldots$ sorted with respect to the corresponding eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots$ solve the FPCA problem (1.5). The following link between eigenvalues and eigenfunctions holds:

$$\lambda_m = \operatorname{Var}(f_m) = \operatorname{Var}\left[\int \gamma_m(t) X(t) dt \right] = \int \int \gamma_m(s) \operatorname{Cov}(s,t) \gamma_m(t) ds dt.$$

In the sampling problem, the unknown covariance function $\operatorname{Cov}(s,t)$ needs to be replaced by the sample covariance function $\widehat{\operatorname{Cov}}(s,t)$. Dauxois, Pousse, and Romain (1982) show that the eigenfunctions and eigenvalues are consistent estimators for λ_m and γ_m and derive some asymptotic results for these estimators.

1.4.1 Basis expansion

Suppose that the weight function γ has expansion

$$\gamma = \sum_{l=1}^{L} \mathbf{b}_l \Theta_l(t) = \boldsymbol{\Theta}^\top \mathbf{b}.$$

Using this notation we can rewrite the left hand side of eigenequation (1.6):

$$\begin{aligned} \int \operatorname{Cov}(s,t) \gamma(t) dt &= \int \boldsymbol{\Theta}(s)^\top \operatorname{Cov}(\mathbf{C}) \boldsymbol{\Theta}(t) \boldsymbol{\Theta}(t)^\top \mathbf{b} dt \\ &= \boldsymbol{\Theta}^\top \operatorname{Cov}(\mathbf{C}) \mathbf{W} \mathbf{b}, \end{aligned}$$

so that:

$$\text{Cov}(\mathbf{C})\mathbf{W}\mathbf{b} = \lambda\mathbf{b}.$$

The functional scalar product $\langle\gamma_l, \gamma_k\rangle$ corresponds to $\mathbf{b}_l^\top\mathbf{W}\mathbf{b}_k$ in the truncated basis framework, in the sense that if two functions γ_l and γ_k are orthogonal, the corresponding coefficient vectors $\mathbf{b}_l, \mathbf{b}_k$ satisfy $\mathbf{b}_l^\top\mathbf{W}\mathbf{b}_k = 0$. Matrix $\mathbf{W}$ is symmetric by definition. Thus, defining $\mathbf{u} = \mathbf{W}^{1/2}\mathbf{b}$, one needs to solve finally a symmetric eigenvalue problem:

$$\mathbf{W}^{1/2}\text{Cov}(\mathbf{C})\mathbf{W}^{1/2}\mathbf{u} = \lambda\mathbf{u},$$

and to compute the inverse transformation $\mathbf{b} = \mathbf{W}^{-1/2}\mathbf{u}$. For the orthonormal functional basis (i.e. also for the Fourier basis) $\mathbf{W} = \mathbf{I}$, i.e. the problem of FPCA is reduced to the multivariate PCA performed on the matrix $\mathbf{C}$.

Using the FPCA method on the IV-strings for 1M and 2M maturities we obtain the eigenfunctions plotted in Figure 1.4. It can be seen, that the eigenfunctions are too rough. Intuitively, this roughness is caused by the flexibility of the functional basis. In the next section we present a way of incorporating the smoothing directly into the PCA problem.

1.5 Smoothed principal components analysis

As we can see in Figure 1.4, the resulting eigenfunctions are often very rough. Smoothing them could result in a more natural interpretation of the obtained weight functions. Here we apply a popular approach known as roughness penalty. The downside of this technique is that we loose orthogonality in the L^2 sense.

Assume that the underlying eigenfunctions of the covariance operator have a continuous and square-integrable second derivative. Let $\mathcal{D}\gamma = \gamma'(t)$ be the first derivative operator and define the roughness penalty by $\Psi(\gamma) = ||\mathcal{D}^2\gamma||^2$. Moreover, suppose that γ_m has square-integrable derivatives up to degree four and that the second and the third derivatives satisfy one of the following conditions:

1. $\mathcal{D}^2\gamma$, $\mathcal{D}^3\gamma$ are zero at the ends of the interval J,
2. the periodicity boundary conditions of γ,$\mathcal{D}\gamma$, $\mathcal{D}^2\gamma$, and $\mathcal{D}^3\gamma$ on J.

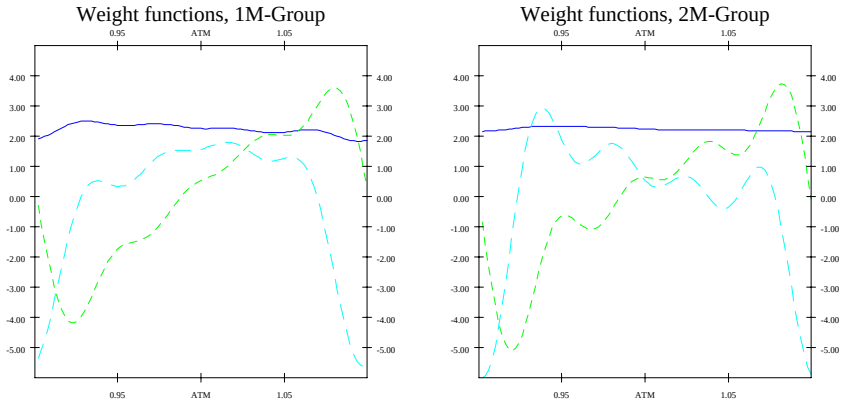

Figure 1.4: Weight functions for $1M$ and $2M$ maturity groups. Blue solid lines, $\hat{\gamma}_1^{1M}$ and $\hat{\gamma}_1^{2M}$, are the first eigenfunctions, green finely dashed lines, $\hat{\gamma}_2^{1M}$ and $\hat{\gamma}_2^{2M}$, are the second eigenfunctions, and cyan dashed lines, $\hat{\gamma}_3^{1M}$ and $\hat{\gamma}_3^{2M}$, are the third eigenfunctions.

STFfda04.xpl

Then we can rewrite the roughness penalty in the following way:

$$\begin{aligned} ||\mathcal{D}^2\gamma||^2 &= \int \mathcal{D}^2\gamma(s)\mathcal{D}^2\gamma(s)ds \\ &= \mathcal{D}\gamma(u)\mathcal{D}^2\gamma(u) - \mathcal{D}\gamma(d)\mathcal{D}^2\gamma(d) - \int \mathcal{D}\gamma(s)\mathcal{D}^3\gamma(s)ds \quad (1.7) \\ &= \gamma(u)\mathcal{D}^3\gamma(u) - \gamma(d)\mathcal{D}^3\gamma(d) - \int \gamma(s)\mathcal{D}^4\gamma(s)ds \quad (1.8) \\ &= \langle\gamma, \mathcal{D}^4\gamma\rangle, \quad (1.9) \end{aligned}$$

where d and u are the boundaries of the interval J and the first two elements in (1.7) and (1.8) are both zero under any of the conditions mentioned above.

Given a eigenfunction γ with norm $||\gamma||^2 = 1$, we can penalize the sample variance of the principal component by dividing it by $1 + \alpha\langle\gamma, \mathcal{D}^4\gamma\rangle$:

$$PCAPV \stackrel{\text{def}}{=} \frac{\int\int \gamma(s)\widehat{\text{Cov}}(s,t)\gamma(t)dsdt}{\int \gamma(t)(\mathcal{I} + \alpha\mathcal{D}^4)\gamma(t)dt}, \quad (1.10)$$

where $\mathcal{I}$ denotes the identity operator. The maximum of the penalized sample variance (PCAPV) is an eigenfunction γ corresponding to the largest eigenvalue of the generalized eigenequation:

$$\int \widehat{\mathrm{Cov}}(s,t)\gamma(t)dt = \lambda(\mathcal{I} + \alpha\mathcal{D}^4)\gamma(s). \tag{1.11}$$

As already mentioned above, the resulting weight functions (eigenfunctions) are no longer orthonormal in the L^2 sense. Since the weight functions are used as smoothed estimators of principal components functions, we need to rescale them to satisfy $||\gamma_l||^2 = 1$. The weight functions γ_l can be also interpreted as orthogonal in the modified scalar product of the Sobolev type

$$(f,g) \stackrel{\mathrm{def}}{=} \langle f,g\rangle + \alpha\langle \mathcal{D}^2 f, \mathcal{D}^2 g\rangle.$$

A more extended theoretical discussion can be found in Silverman (1991).

1.5.1 Basis expansion

Define $\mathbf{K}$ to be a matrix whose elements are $\langle \mathcal{D}^2\Theta_j, \mathcal{D}^2\Theta_k\rangle$. Then the generalized eigenequation (1.11) can be transformed to:

$$\mathbf{W}\,\mathrm{Cov}(\mathbf{C})\mathbf{W}\mathbf{u} = \lambda(\mathbf{W} + \alpha\mathbf{K})\mathbf{u}. \tag{1.12}$$

Using Cholesky factorization $\mathbf{L}\mathbf{L}^\top = \mathbf{W} + \alpha\mathbf{K}$ and defining $\mathbf{S} = \mathbf{L}^{-1}$ we can rewrite (1.12) as:

$$\{\mathbf{S}\mathbf{W}\,\mathrm{Cov}(\mathbf{C})\mathbf{W}\mathbf{S}^\top\}(\mathbf{L}^\top\mathbf{u}) = \lambda\mathbf{L}^\top\mathbf{u}.$$

Applying Smoothed Functional PCA (SPCA) to the IV-strings, we get the smooth-eigenfunctions plotted in Figure 1.5. We use $\alpha = 10^{-7}$, the aim is to use a rather small degree of smoothing, in order to replace the high frequency fluctuations only. Some popular methods, like cross-validation, could be employed as well, Ramsay and Silverman (1997).

The interpretation of the weight functions displayed in Figure 1.5 is as follows: The first weight function (solid blue) represents clearly the level of the volatility – weights are almost constant and positive. The second weight function (finely dashed green) changes sign near the at-the-money point, i.e. can be interpreted as the in-the-money/out-of-the-money identification factor or slope.

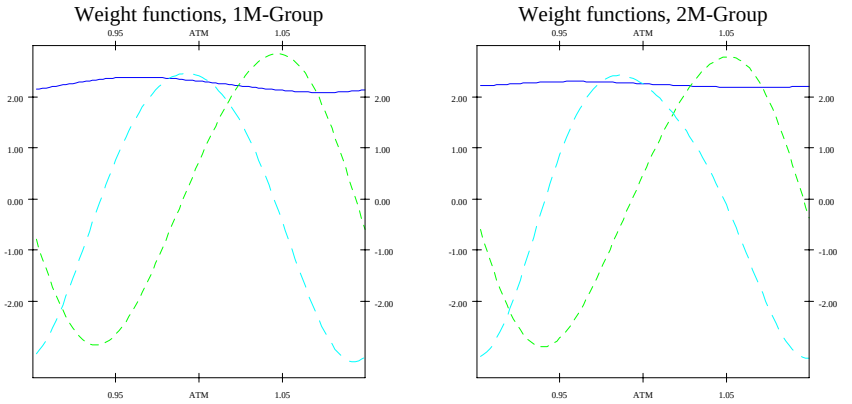

Figure 1.5: Smoothed weight functions with $\alpha = 10^{-7}$. Blue solid lines, $\hat{\gamma}_1^{1M}$ and $\hat{\gamma}_1^{2M}$, are the first eigenfunctions, green finely dashed lines, $\hat{\gamma}_2^{1M}$ and $\hat{\gamma}_2^{2M}$, are the second eigenfunctions, and cyan dashed lines, $\hat{\gamma}_3^{1M}$ and $\hat{\gamma}_3^{2M}$, are the third eigenfunctions.

STFfda05.xpl

The third (dashed cyan) weight function may play the part of the measure for a deep in-the-money or out-of-the-money factor or curvature. It can be seen that the weight functions for the 1M $(\widetilde{\gamma}_1^{1M}, \widetilde{\gamma}_2^{1M}, \widetilde{\gamma}_3^{1M})$ and 2M maturities $(\widetilde{\gamma}_1^{2M}, \widetilde{\gamma}_2^{2M}, \widetilde{\gamma}_3^{2M})$ have a similar structure. From a practical point of view it can be interesting to try to get common estimated eigenfunctions (factors in the further analysis) for both groups. In the next section, we introduce the estimation motivated by the Common Principal Component Model.

1.6 Common principal components model

The Common Principal Components model (CPC) in the multivariate setting can be motivated as the model for similarity of the covariance matrices in the k-sample problem, Flury (1988). Having k random vectors, $\mathbf{x}_{(1)}, \mathbf{x}_{(2)}, \ldots, \mathbf{x}_{(k)} \in$

$\mathbb{R}^p$ the CPC-Model can be written as:

$$\mathbf{\Psi}_j \stackrel{\text{def}}{=} \text{Cov}(\mathbf{x}_{(j)}) = \mathbf{\Gamma}\mathbf{\Lambda_j}\mathbf{\Gamma}^\top,$$

where $\mathbf{\Gamma}$ is an orthogonal matrix and $\mathbf{\Lambda}_j = \text{diag}(\lambda_{i1}, \ldots, \lambda_{ip})$. This means that eigenvectors are the same across samples and just the eigenvalues – variances of the principal component scores (1.4) differ.

Using the normality assumption, the sample covariance matrices $\mathbf{S}_j, j = 1, \ldots, k$, are Wishart-distributed:

$$\mathbf{S}_j \sim W_p(n_j, \mathbf{\Psi}_j/n_j),$$

and the CPC model can be estimated using maximum likelihood estimation with likelihood-function:

$$\text{L}(\mathbf{\Psi}_1, \mathbf{\Psi}_2, \ldots, \mathbf{\Psi}_k) = C \prod_{j=1}^{k} \exp\left\{\text{tr}\left(-\frac{n_j}{2}\mathbf{\Psi}_j^{-1}\mathbf{S}_j\right)\right\} (\det\mathbf{\Psi_j})^{-n_j/2}.$$

Here C is a factor that does not depend on the parameters and n_j is the number of observations in group j. The maximization of this likelihood function is equivalent to:

$$\prod_{j=1}^{k} \left\{\frac{\det \text{ diag}(\mathbf{\Gamma}^\top\mathbf{S_j}\mathbf{\Gamma})}{\det(\mathbf{\Gamma}^\top\mathbf{S_j}\mathbf{\Gamma})}\right\}^{n_j}, \tag{1.13}$$

and the maximization of this criterion is performed by the so-called Flury-Gautschi(FG)-algorithm, Flury (1988).

As shown in Section 1.4, using the functional basis expansion, the FPCA and SPCA are basically implemented via the spectral decomposition of the "weighted" covariance matrix of the coefficients. In view of the minimization property of the FG algorithm, the diagonalization procedure optimizing the criterion (1.13) can be employed. However, the obtained estimates may not be maximum likelihood estimates.

Using this procedure for the IV-strings of 1M and 2M maturity we get "common" smoothed eigenfunctions. The first three common eigenfunctions ($\widetilde{\gamma}_1^c, \widetilde{\gamma}_2^c$, $\widetilde{\gamma}_3^c$) are displayed in Figures 1.6–1.8. The solid blue curve represents the estimated eigenfunction for the 1M maturity, the finely dashed green curve for the 2M maturity and the dashed black curve is the common eigenfunction estimated by the FG-algorithm.

Assuming that $\hat{\sigma}_i(\kappa, \tau)$ are centered for $\tau = 1M$ and $2M$ (we subtract the sample mean of corresponding group from the estimated functions), we may

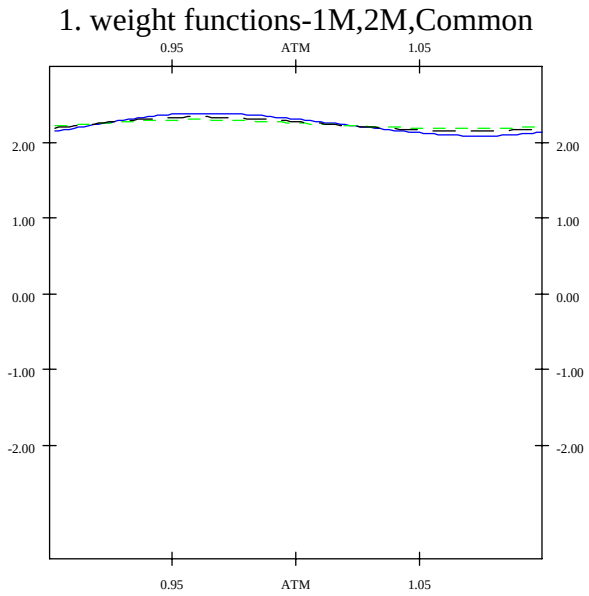

Figure 1.6: First weight functions, $\alpha = 10^{-7}$, solid blue line is the weight function of the 1M maturity group ($\hat{\gamma}_1^{1M}$), finely dashed green line of the 2M maturity group ($\hat{\gamma}_1^{2M}$), and dashed black line is the common eigenfunction ($\widetilde{\gamma}_1^c$), estimated from both groups.

use the obtained weight functions in the factor model of the IV dynamics of the form:

$$\widetilde{\sigma}_i(\kappa,\tau) = \sum_{j=1}^{R} \widetilde{\gamma}_j^c(\kappa)\langle\widetilde{\gamma}_j^c(\kappa), \hat{\sigma}_i(\kappa,\tau)\rangle, \tag{1.14}$$

where $\tau \in \{1M, 2M\}$ and R is the number of factors. Thus $\widetilde{\sigma}_i$ is an alternative estimation of σ_i. This factor model can be used for simulation applications like Monte Carlo VaR. Especially the usage of Common Principal Components $\widetilde{\gamma}_j^c(\kappa)$ reduces the high-dimensional IV-surface problem to a small number of functional factors.

In addition, an econometric approach, successfully employed by Fengler, Härdle, and Mammen (2004) can be employed. It consists of fitting an appropri-

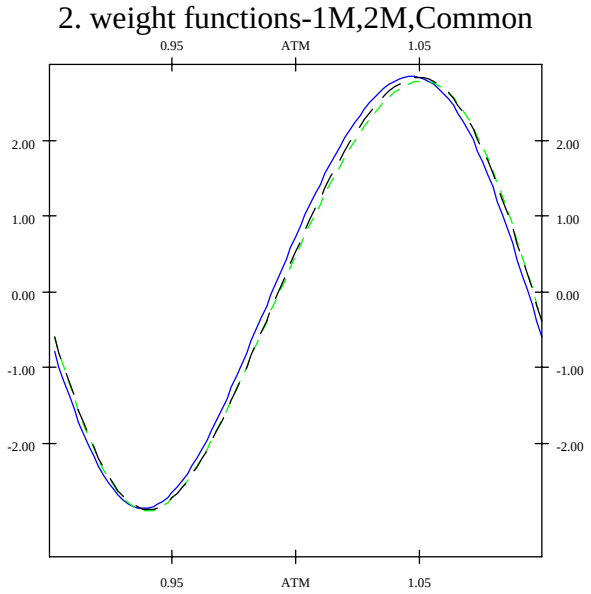

Figure 1.7: Second eigenfunctions, $\alpha = 10^{-7}$, solid blue line is the weight function of the 1M maturity group ($\hat{\gamma}_2^{1M}$), finely dashed green line of the 2M maturity group ($\hat{\gamma}_2^{2M}$), and dashed black line is the common eigenfunction ($\widetilde{\gamma}_2^c$), estimated from both groups.

ate model to the time series of the estimated principal component scores, $\widetilde{f}_{ij}^c(\tau) = \langle \widetilde{\gamma}_j^c(\kappa),\ \hat{\sigma}_i(\kappa,\tau) \rangle$, as displayed in Figure 1.9. Note that $\hat{\sigma}_i(\kappa,\tau)$ are centered again (sample means are zero). The fitted time series model can be used for forecasting future IV functions.

There are still some open questions related to this topic. First of all, the practitioner would be interested in a good automated choice of the parameters of our method (dimension of the truncated functional basis L and smoothing parameter α). The application of the Fourier coefficients in this framework seems to be reasonable for the volatility smiles (U-shaped strings), however for the volatility smirks (typically monotonically decreasing strings) the performance is rather bad. In particular, the variance of our functional objects and the shape of our weight functions at the boundaries is affected. The application

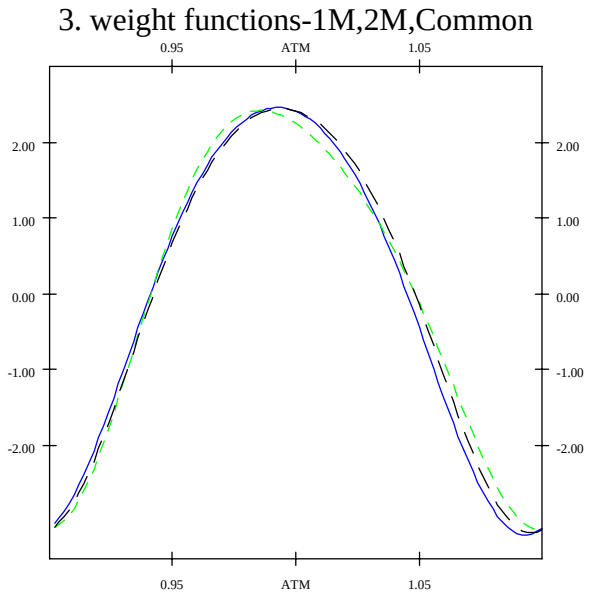

Figure 1.8: Third eigenfunctions, $\alpha = 10^{-7}$, solid blue line is the weight function of the 1M maturity group ($\hat{\gamma}_3^{1M}$), finely dashed green line of the 2M maturity group ($\hat{\gamma}_3^{2M}$), and dashed black line is the common eigenfunction ($\widetilde{\gamma}_3^{c}$), estimated from both groups.

of regression splines in this setting seems to be promising, but it increases the number of smoothing parameters by the number and the choice of the knots – problems which are not generally easy to deal with. The next natural question, which is still open concerns the statistical properties of the technique and the testing procedure for the Functional Common PCA model. Finally, using the data for a longer time period one may also analyze the longer maturities like 3 months or 6 months.

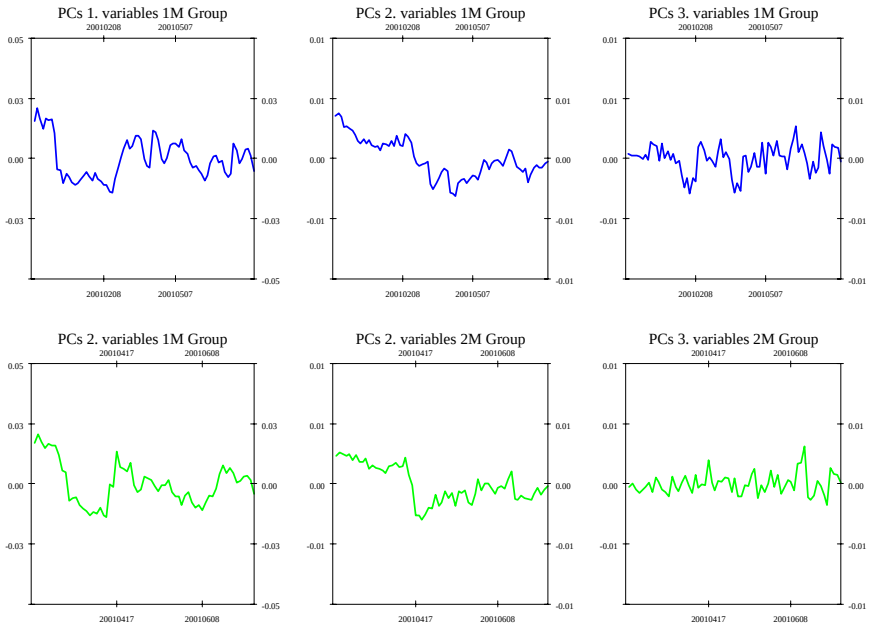

Figure 1.9: Estimated principal component scores, $\widetilde{f}^c_{i1}(1M)$, $\widetilde{f}^c_{i2}(1M)$, and $\widetilde{f}^c_{i3}(1M)$ for 1M maturity – first row, and $\widetilde{f}^c_{i1}(2M)$, $\widetilde{f}^c_{i2}(2M)$, and $\widetilde{f}^c_{i3}(2M)$ for 2M maturity – second row; $\alpha = 10^{-7}$.

Bibliography

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities, *Journal of Political Economy*, **81**: 637:654.

Dauxois, J., Pousse, A., and Romain, Y. (1982). Asymptotic Theory for the Principal Component Analysis of a Vector Random Function: Some Applications to Statistical Inference, *Journal of Multivariate Analysis* **12**: 136-154.

Flury, B. (1988). *Common Principal Components and Related Models*, Wiley, New York.

Fengler, M., Härdle, W., and Schmidt, P. (2002). Common Factors Governing VDAX Movements and the Maximum Loss, *Journal of Financial Markets and Portfolio Management* **16(1)**: 16-29.

Fengler, M., Härdle, W., and Villa, P. (2003). The Dynamics of Implied Volatilities: A common principle components approach, *Review of Derivative Research* **6**: 179-202.

Fengler, M., Härdle, W., and Mammen, E. (2004). Implied Volatility String Dynamics, CASE Discussion Paper, http://www.case.hu-berlin.de.

Föllmer, H. and Schied A. (2002). *Stochastic Finance*, Walter de Gruyter.

Härdle, W. (1990). *Applied Nonparametric Regression*, Cambridge University Press.

Hafner, R. and Wallmeier, M. (2001). The Dynamics of DAX Implied Volatilities, *International Quarterly Journal of Finance* **1(1)**: 1-27.

Härdle, W. and Simar, L. (2003). *Applied Multivariate Statistical Analysis*, Springer-Verlag Berlin Heidelberg.

Kneip, A. and Utikal, K. (2001). Inference for Density Families Using Functional Principal Components Analysis, *Journal of the American Statistical Association* **96**: 519-531.

Ramsay, J. and Silverman, B. (1997). *Functional Data Analysis*, Springer, New York.

Rice, J. and Silverman, B. (1991). Estimating the Mean and Covariance Structure Nonparametrically when the Data are Curves, *Journal of Royal Statistical Society, Series B* **53**: 233-243.

Silverman, B. (1996). Smoothed Functional Principal Components Analysis by Choice of Norm, *Annals of Statistics* **24**: 1-24.

SFB 649 Discussion Paper Series

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

001 "Nonparametric Risk Management with Generalized Hyperbolic Distributions" by Ying Chen, Wolfgang Härdle and Seok-Oh Jeong, January 2005.
002 "Selecting Comparables for the Valuation of the European Firms" by Ingolf Dittmann and Christian Weiner, February 2005.
003 "Competitive Risk Sharing Contracts with One-sided Commitment" by Dirk Krueger and Harald Uhlig, February 2005.
004 "Value-at-Risk Calculations with Time Varying Copulae" by Enzo Giacomini and Wolfgang Härdle, February 2005.
005 "An Optimal Stopping Problem in a Diffusion-type Model with Delay" by Pavel V. Gapeev and Markus Reiß, February 2005.
006 "Conditional and Dynamic Convex Risk Measures" by Kai Detlefsen and Giacomo Scandolo, February 2005.
007 "Implied Trinomial Trees" by Pavel Čížek and Karel Komorád, February 2005.
008 "Stable Distributions" by Szymon Borak, Wolfgang Härdle and Rafal Weron, February 2005.
009 "Predicting Bankruptcy with Support Vector Machines" by Wolfgang Härdle, Rouslan A. Moro and Dorothea Schäfer, February 2005.
010 "Working with the XQC" by Wolfgang Härdle and Heiko Lehmann, February 2005.
011 "FFT Based Option Pricing" by Szymon Borak, Kai Detlefsen and Wolfgang Härdle, February 2005.
012 "Common Functional Implied Volatility Analysis" by Michal Benko and Wolfgang Härdle, February 2005.

SFB 649, Spandauer Straße 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".