Approximate Transition Density Estimation of the Stochastic Cusp Model

Jan Voříšek [1]

Abstract. Stochastic cusp model is defined by stochastic differential equation with cubic drift. Its stationary density belongs to the class of generalized normal distributions, allows for skewness, different tail shapes and bimodality. There are two stable equilibria in bimodality case and movement from one equilibrium to another is interpreted as a crash. Qualitative properties of the cusp model were employed to model crashes on financial markets, however, practical applications of the model employed the stationary distribution, which does not take into account the serial dependence between observations.

Because closed-form solution of the transition density is not known, one has to use approximate technique to estimate transition density. This paper extends approximate maximum likelihood method, which relies on the closed-form expansion of the transition density, to incorporate time-varying parameters of the drift function to be driven by market fundamentals. A measure to predict endogenous crashes of the model is proposed using transition density estimates.

Empirical example estimates Iceland Krona depreciation with respect to the British Pound in the year 2001 using differential of interbank interest rates as a market fundamental.
Keywords: multimodal distributions, stochastic cusp model, approximate transition density.

JEL classification: C46
AMS classification: 62F10

1 Introduction

Stationary density of stochastic cusp model belongs to the class of generalized normal distributions. Since it has four parameters, its flexible enough to allow for skewness, kurtosis and bimodality (Cobb et al. [4]). Modes of stationary density correspond to the stable equilibria of differential equation with cubic polynomial. In bimodality case there are two stable equilibria attracting the process and one unstable one between them, which repulses the process and corresponds to the antimode of stationary density. Movement from one stable equilibrium to another, which is defined by switching from one side of unstable equilibrium to another, can be viewed as an intrinsic crash in the system. This potentially large shift towards another stable equilibrium level is considerable advantage of the cusp model in describing certain systems over traditionally used mean reverting linear models, which has just one stable attracting equilibrium. Another advantage is a random walk behavior (under certain parametrization) in the middle of the domain, which was found appropriate for interest rate modeling by Aït-Sahalia [1].

However, complexity of cusp model brings some disadvantages. The major disadvantage lies in non-existence of closed-form solution of the transition density. There are several different approaches to overcome this obstacle: Euler approximation, simulation based methods, binomial approximations, numerical solution of Kolmogorow equations, and Hermite expansions. The last method, proposed by Aït-Sahalia [2] and [3], gives unlike the other methods closed-form approximation of the transition density, that converges to the true likelihood function. This methods allows to consistently estimate parameters of the diffusion processes.

[1] Academy of Sciences of the Czech Republic, Institute of Information Theory and Automation, Pod Vodárenskou věží 4, CZ-182 08 Prague 8, Czech Republic
Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University Prague, Sokolovská 83, CZ-186 75 Prague 8, Czech Republic, vorisek@karlin.mff.cuni.cz

Approach, which utilizes time-varying parameters, but only stationary density of stochastic cusp model were theoretically derived by Creedy and Martin [5] and empirically tested by the same authors [6] for the USD/GBP exchange rate. There were attempts to model currency crises by stationary cusp density (e.g. [7], [10]), some of them neglecting serial dependence completely, some reflecting it in variability of estimates. This paper contributes to overcome the issue of time dependence by extension of approximate transition density method to include exogenous variables, which drive parameters of the model.

The rest of this paper is organized as follows. Section 2 recall a method, which approximates transition density function. Estimation of stochastic cusp model with time-varying parameters and a measure to access potential of crash is introduced in Section 3. Section 4 utilizes proposed methodology on real data example of ISK/GBP exchange rate using daily data for the period from July 1, 1999 to December 31, 2004. Section 5 summarizes the results and concludes.

2 Approximate transition density function

Transition density of a general diffusion process

$$dX_t = \mu(X_t;\theta)dt + \sigma(X_t;\theta)dW_t, \tag{1}$$

does not have an analytical form. There are different techniques, how it can be approximated including Euler approximation, simulation-based methods or a numerical solution of forward Kolmogorov equation (see [8]). Aït-Sahalia ([3]) proposed a method based on Hermite polynomials to derive closed-form expansion for the transition density. Assume following *assumptions* :

1. The functions $\mu(x;\theta)$ and $\sigma(x;\theta)$ are infinitely differentiable in x, and three times continuously differentiable in θ, for all $x \in D_X = (-\infty,\infty)$ and $\theta \in \Theta$.
2. There exists a constant ζ such that $\sigma(x;\theta) > \zeta > 0$ for all $x \in D_X$ and $\theta \in \Theta$.
3. For all $\theta \in \Theta$, $\mu(x;\theta)$ and its derivatives with respect to x and θ have at most polynomial growth near the boundaries and
$$\lim_{x\to\pm\infty} -\frac{1}{2}\left(\mu^2(x;\theta) + \frac{\partial\mu(x;\theta)}{\partial x}\right) < \infty,$$

and a *transformation* of original process X into a new process Y as

$$Y \equiv \gamma(X;\theta) = \int^X \frac{1}{\sigma(u;\theta)}\mathrm{d}u.$$

By applying Itô's Lemma, Y has unit diffusion

$$\mathrm{d}Y_t = \mu_Y(Y_t;\theta)\mathrm{d}t + \mathrm{d}W_t,$$

where the drift is given by

$$\mu_Y(y;\theta) = \frac{\mu(\gamma^{-1}(y;\theta);\theta)}{\sigma(\gamma^{-1}(y;\theta);\theta)} - \frac{1}{2}\frac{\partial\sigma(x;\theta)}{\partial x}|_{x=\gamma^{-1}(y;\theta)}.$$

Than approximation of the log-transition density of y given initial values y_0 and a time step Δ is given by:

$$l_Y^{(K)}(y|y_0,\Delta;\theta) = -\frac{1}{2}\log(2\pi\Delta) - \frac{C^{(-1)}(y|y_0;\theta)}{\Delta} + \sum_{k=0}^{K} C^{(k)}(y|y_0;\theta)\frac{\Delta^k}{k!} \tag{2}$$

where K is order of expansion connected to the power of Δ and coefficients $C^{(k)}(y|y_0;\theta)$ can be calculated by substitution of proposed solution (2) into forward and backward Kolmogorov equations. Given $l_Y^{(K)}$, the expression for $l_X^{(K)}$ is given by the Jacobian formula

$$l_X^{(K)}(x|x_0,\Delta) = -\frac{1}{2}\log(2\pi\Delta\sigma^2(x)) - \frac{(\gamma(x)-\gamma(x_0))^2}{2\Delta} + \sum_{k=0}^{K} C^{(k)}(\gamma(x)|\gamma(x_0);\theta)\frac{\Delta^k}{k!}. \quad (3)$$

3 Stochastic model cusp

The univariate stochastic cusp model can be characterized by nonlinear diffusion process for the variable of interest

$$dX_t = \left(\alpha + \beta\,\frac{X_t-\lambda}{\sigma} - \left(\frac{X_t-\lambda}{\sigma}\right)^3\right)\frac{1}{2\sigma}dt + \sigma_w dW_t, \quad (4)$$

where W_t is standard Brownian motion, $\sigma > 0$ and $\sigma_w > 0$. Stationary density can be expressed analytically

$$f_s(x;\theta) = \eta_s(\theta)\exp\left[\alpha\,\frac{x-\lambda}{\sigma} + \frac{\beta}{2}\left(\frac{x-\lambda}{\sigma}\right)^2 - \frac{1}{4}\left(\frac{x-\lambda}{\sigma}\right)^4\right], \quad (5)$$

where $\eta_s(\theta)$ is normalizing constant.

For identifying bimodality of the stationary probability density function (5) serves statistic called Cardan's discriminant

$$\delta_C = \left(\frac{\alpha}{2}\right)^2 - \left(\frac{\beta}{3}\right)^3. \quad (6)$$

The parameters α (*asymetry*) and β (*bifurcation*) are invariant with respect to changes in λ (*location*) and σ (*scale*), as is δ_C, and they have following approximate interpretations [4]. If $\delta_C \geq 0$ then α measures skewness and β kurtosis, while $\delta_C < 0$ then α indicates the relative height of the two modes and β their relative separations.

3.1 Time-varying parameters

Following for example Creedy and Martin [6] or Fernandes [7], one can allow parameters α and β from (4) to be time varying:

$$dx_t = \left(\alpha(\xi_t) + \beta(\xi_t)\,\frac{x_t-\lambda}{\sigma} - \left(\frac{x_t-\lambda}{\sigma}\right)^3\right)\frac{1}{2\sigma}dt + \sigma_w dW_t, \quad (7)$$

where ξ_t is a vector of market fundamentals that are strictly exogenous with respect to x_t. Using approximate transition density (3) log-likelihood of the observed values x_i, ξ_i is given by:

$$ll^{(2)} = \sum_{i=1}^{n-1}\left[-\frac{(z_{i+1}-z_i)^2}{2\Delta_\sigma} + \frac{\alpha_i}{2}\frac{z_{i+1}-z_i}{\sigma_w} + \frac{\beta_i}{4}\frac{z_{i+1}^2-z_i^2}{\sigma_w} - \frac{z_{i+1}^4-z_i^4}{8\sigma_w} + \right.$$

$$+\frac{\Delta_\sigma}{\sigma_w^2}\left(-\frac{\alpha_i^2}{8} + \frac{\alpha_i}{16}\left(\kappa_i^3 - 2\beta\kappa_i\right) - \frac{\beta_i^2}{24}\kappa_i^2 + \frac{\beta_i}{20}\left(-5\sigma_w + \kappa_i^4\right) + \frac{\sigma_w}{4}\kappa_i^2 - \frac{1}{56}\kappa_i^6\right) +$$

$$\left. +\frac{\Delta_\sigma^2}{\sigma_w^2}\left(\frac{\sigma_w}{8} - \frac{\beta_i^2}{48} + \frac{\beta_i}{40}\kappa_i^2(3,4) + \frac{\alpha_i}{16}\kappa_i - \frac{1}{112}\kappa_i^4(5,8,9)\right) - \frac{1}{2}\log(2\pi\Delta\sigma_w^2)\right] \quad (8)$$

where $\Delta_\sigma = \Delta/\sigma^2$, $\alpha_i = \alpha_0 + \alpha_1\xi_i$, $\beta_i = \beta_0 + \beta_1\xi_i$, $z_i = (x_i/\sigma_w - \lambda)/\sigma$ and $\kappa_i^j(c) = \sum_{k=0}^{j} c_k\ z_{i+1}^k z_i^{j-k}$.

By differentiation of log-likelihood (8) with respect to α_0 and β_0 and laying down to zero, one can express these parameters wrt. other parameters

$$\hat{\alpha}_0 = \frac{2c^0_{13}c^0_{21} - c^0_{12}c^0_{22} + 2\beta_1(c^1_{22}c^0_{13} - c^0_{22}c^1_{13}) + \alpha_1(4c^0_{13}c^1_{31} - c^0_{22}c^1_{22})}{(c^0_{22})^2 - 4c^0_{13}c^0_{31}}, \tag{9}$$

$$\hat{\beta}_0 = \frac{2c^0_{12}c^0_{31} - c^0_{21}c^0_{22} + 2\alpha_1(c^1_{22}c^0_{31} - c^0_{22}c^1_{31}) + \beta_1(4c^0_{31}c^1_{13} - c^0_{22}c^1_{22})}{(c^0_{22})^2 - 4c^0_{13}c^0_{31}}, \tag{10}$$

where coefficients c_{ij} are as follows:

$$\begin{aligned}
c^k_{12} &= \textstyle\sum_{i=1}^{n-1} \xi^k_i \left(\frac{1}{4}(z^2_{i+1} - z^2_i) + \Delta_\sigma \left(-\frac{1}{4} + \frac{1}{20}\kappa^4_i\right) + \Delta^2_\sigma \frac{\kappa^2_i(3,4)}{40}\right), \\
c^k_{13} &= \textstyle\sum_{i=1}^{n-1} \xi^k_i \left(-\Delta_\sigma \frac{\kappa^2_i}{24} - \Delta^2_\sigma \frac{1}{48}\right), \\
c^k_{21} &= \textstyle\sum_{i=1}^{n-1} \xi^k_i \left(\frac{1}{2}(z_{i+1} - z_i) + \Delta_\sigma \frac{\kappa^3_i}{16} + \Delta^2_\sigma \frac{\kappa_i}{16}\right), \\
c^k_{22} &= -\textstyle\sum_{i=1}^{n-1} \xi^k_i \Delta_\sigma \frac{\kappa_i}{8}, \\
c^k_{31} &= -\textstyle\sum_{i=1}^{n-1} \xi^k_i \Delta_\sigma \frac{1}{8}.
\end{aligned}$$

Besides, one can express parameters α_1 and β_1 in a similar way to further reduce parameter space, which means that only parameters λ, σ and σ^2_w needs to be estimated numerically. Employing of the transition density estimation allows to estimate variance of dW_t in contrast with stationary density estimation.

3.2 Probability of crash

Denote probability of extreme event by

$$\pi(x|x_0, \Delta; \theta, c) = \int_{-\infty}^{c} \exp[\, l(x|x_0, \Delta; \theta)]dx, \tag{11}$$

which for c corresponding to the antimode of the stationary density in bimodality case represents probability of switching to lower stable equilibria from starting point $x_0 > c$. If Cardan's discriminant (6) is lower then zero, than antimode c_a can be calculated as:

$$c_a = -2\ \mathrm{sign}(a)\sqrt{\frac{b}{3}}\ \cos\left(\frac{1}{3}\arctan\left(\frac{2}{\mathrm{abs}(a)}\sqrt{-\frac{a^2}{4} + \frac{b^3}{27}}\right) + \arctan\left(\sqrt{3}\right)\right). \tag{12}$$

In general, $\pi(x|x_0, \Delta; \theta, c)$ can be used with an arbitrary c, when it represents standard risk measure approach. It measures large changes in unimodality case as well, however they occur exclusively due to the stochasticity of the process.

4 Empirical example

Following [10] and [7], the example for illustration of proposed methodology involves daily observed exchange rate between Icelandic Krona and British Pound in the period from July 1, 1999 to December 31, 2004, where as a market fundamental is used differential of one month interbank interest rate of the corresponding country[1]. The basic statistical characteristic is presented in Table 1. The interest rate differential is a commonly used macroeconomic indicator of the soundness of the banking sector as high interest rate differentials often precede increases in nonperforming loans (see [9]).

Keeping the modeled quantities as usual (e.g. [6], [10]), we will use the logarithm of the exchange rate x_t and logarithm of the interest rate differential ξ_t, where variable parameters of drift (7) are

[1]Data from the Central Bank of Iceland and Bank of England.

	ISK/GBP	r^I/r^B
mean	128.97	1.0397
median	127.55	1.0421
std. dev.	9.94	0.0222
skewness	0.66	0.2694
kurtosis	2.71	1.8909
minimum	112.75	1.0076
maximum	156.3	1.0886

Table 1 Descriptive statistics of exchange rate and interest rate differential.

given by $\alpha_t = \alpha_0 + \alpha_1 \xi_{t-1}$ and $\beta_t = \beta_0 + \beta_1 \xi_{t-1}$. Table 2 discloses the results of numerical maximum likelihood estimation of the diffusion process. It reveals that the stationary and transition estimates of parameters are similar, but the errors of estimates are remarkably higher for transition density estimation. There are two possible explanation for that: firstly the non-normality of estimates, but this holds for stationary density estimation as well, second explanation holds just for the transition model, where the last observation of exchange rate explains substantial part of distribution of next observation. However, Figure 1 shows, that the switching probability (11) indicated possibility of depreciation of the Krona in the first quarter of 2001, before the real depreciation came. Except of that Figure 1 displays evolution of Cardan's discriminant calculated according to both stationary and transition estimates, together with scaled exchange rate, interest rate differential and probability of adverse equilibria in a one-step ahead forecast.

parameter	stationary est.	stat. std. error	transition est.	trans. std. error
α_0	-0.529	0.054	-0.525	1.273
α_1	0.615	0.042	1.047	0.933
β_0	0.021	0.173	-0.054	1.562
β_1	2.312	0.099	2.018	1.328
λ	0.046	0.025	0.031	0.268
σ	0.906	0.018	0.755	0.105
σ_w			1.207	0.023

Table 2 Estimation results for stationary and transition density.

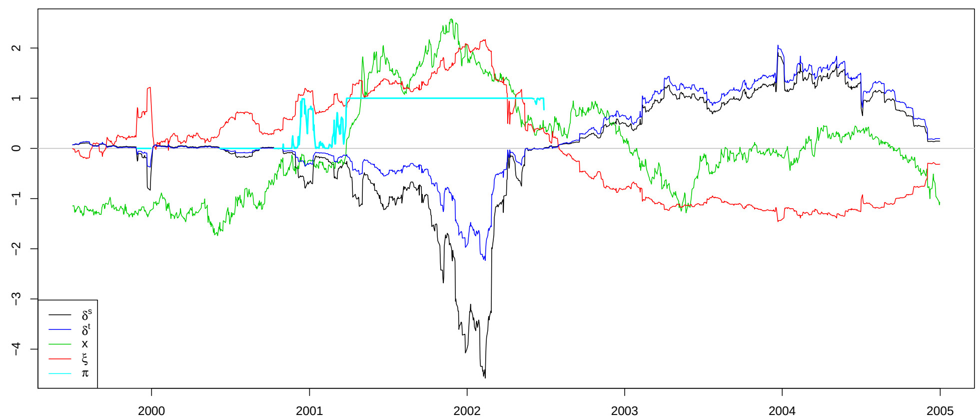

Figure 1 Stationary and transition evolution of Cardan's discriminant, exchange rate, interest rate differential and probability of adverse equilibria.

5 Conclusion

This paper extends approximate maximum likelihood approach to the transition density estimation by time-varying parameters of the stochastic cusp model. It shows how to simplify the estimation and the measure for indicating endogenous crash of the system is introduced.

To illustrate the performance of proposed methodology is used the example concerning Iceland Krona and British Pound exchange rate with interest rate differential as a market fundamental. Comparison with the stationary density approach reveals the behavior of the exchange rate close to the random walk, but the measure of crash was able to indicate the Krona depreciation in advance.

Acknowledgements

Supported by the grant No. P402/12/G097 of the Czech Science Foundations.

References

[1] Aït-Sahalia, Y.: Testing continuous-time models of the spot interest rate. *Review of Financial Studies* **vol. 9** (1996), 385-426.

[2] Aït-Sahalia, Y.: Maximum Likelihood Estimation of Discretely Sampled Diffusion: A Closed-form Approximation Approach. *Econometrica* **vol. 70**, no. 1 (2002), 223–262.

[3] Aït-Sahalia, Y.: Closed-form Likelihood Expansions for Multivariate Diffusions. *The Annals of Statistics* **vol. 36**, no. 2 (2008), 906–937.

[4] Cobb, L., Koppstein, P., and Chen, N. H. : Estimation and Moment Recursion Relations for Multimodal Distributions of the Exponential Family. *Journal of the American Statistical Association vol. 78*, no. 381 (1983), 124–130.

[5] Creedy, J., and Martin, V. : Multiple Equilibria and Hysteresis in Simple Exchange Models. *Economic Modelling* **vol. 10**, Issue 4 (1993), 339–347.

[6] Creedy, J., and Martin, V. : A Non-Linear Model of the Real US/UK Exchange Rate. *Journal of Applied Econometrics* **vol. 11**, no. 6 (1996), 669–686.

[7] Fernandes, M.: Financial crashes as endogenous jumps: estimation, testing and forecasting. *Journal of Economic Dynamics and Control* **vol. 30**, Issue 1 (2006), 111–141.

[8] Jensen, B. and Poulsen, R. : Transition densities of diffusion processes: Numerical comparison of approximation techniques. *Journal of Derivatives* **vol. 9** (2002), 18-32.

[9] Kaminsky, G. L. and Reinhart, C : The twin crises: causes of banking and balance-of-payments crises. *American Economic Review* **vol. 89** (1996), 473–500.

[10] Koh, S. K., Fong, W. M., Chan, F. : A Cardans discriminant approach to predicting currency crashes. *Journal of International Money and Finance* **vol. 26** (2007), 131–148.