Bayesian vector auto-regression model with Laplace errors applied to financial market data

Jan Šindelář[1]

Abstract. The article presents alternative version of Bayesian vector auto-regression model with Laplace distributed innovations. Bayesian estimation in such model is more computationally demanding than estimation in a model with normally distributed innovations, but because of the heavier tails of Laplace distribution, is more robust. In the article I try to present the way of proceeding with the estimation, obtaining a full posterior distribution of the parameters as a result. At the end an efficient algorithm is sketched, but this part of the research is still unfinished.

Keywords: Financial modeling, auto-regression, time series, robust

JEL classification: C11, C51
AMS classification: 62F15, 62F35

1 Introduction

Modeling financial market price is a challenging task. Systematic mathematical modeling has been first employed by Bachelier at the beginning of the 20th century [1]. In the middle of the century several branches of research have emerged in the finance community, two of which I consider the main influences for writing the following text. A branch followed first by Black, Scholes, Merton and others [2, 3, 4] being motivated by the hypotheses of Fama [5] and Samuelson [6] in stating the market to be efficient and the number of transactions in the market to be large in every time interval and a branch followed by many econometricians, modeling the market price using time series models based on ARMA models described in depth by Box and Jenkins [7].

1.1 Modeling market price in continuous time

The former approach leads to the use of the geometric Brownian motion W_t with drift as a model of market price evolution, where the stochastic process modeling the price Y_t is described by a stochastic differential equation

$$\frac{dY_t}{Y_t} = \alpha_t dt + \sigma_t dW_t \tag{1}$$

where the parameters α_t, σ_t are often considered constant in time. Although such a model is a rich one and is based on foundations evolving now for more than half a century, it has certain drawbacks that cause it not to explain many important phenomena observed in the market data time series. For a moment I set aside the problem of efficient market hypothesis (EMH) itself and I focus on the second major drawback. It is the limit used in the derivation of the model from a random walk that except for making the time continuous also considers a large number (in an extreme case infinite) of transactions in every time interval. The Central Limit Theorem (Donsker's Theorem [8]) then collapses the natural distribution of innovations in the random walk to a Wiener process, if we consider the path of the process to be continuous and a few rather natural conditions of the process to be fulfilled.

In an observed time series from a real market measured empirical distributions of differences of logarithmized data $\ln Y_{t+1} - \ln Y_t$ are often very different from a Gaussian distribution on practically all time scales. These distributions are usually heavy-tailed, meaning they may present a higher risk of high price moves than the often assumed normal distribution. As mentioned by Bouchaud and Potters [9], the differences are quite close to being drawn from a Student distribution. In the branch of financial mathematics research following Merton, the mentioned problem is often resolved by assuming a more general class of Levy processes for modeling the time series [10], since the stochastic calculus used by Merton can be extended to cover these as well, but the study of these processes is out of scope of presented text.

1.2 Use of AR processes

The second branch of using a model of ARMA type to describe the stochastic process underlying market price (and generally market data) generation is not based on an underlying economic theory as much as

[1]Department of Adaptive Systems, Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Pod Vodarenskou vezi 4, 18200 Prague, Czech Republic, e-mail: jenik.sindelar@gmail.com

on the richness of the entire class of such models. Since the models are in contradiction with the stronger forms of EMH, they haven't been used so often in the group of financial mathematicians following Merton, although there have been articles written [11, 12] about their financial application and their use in other areas of application is very common. The class is so rich that a discretized version of the model (1) is its member and therefore, if we have a good structure selection method at hand, which often proves to be unrealistic, we can test, if the use of the Brownian motion-like model is adequate. Another advantage of these models is that since they are discrete time, there is no limit forcing the distribution of the innovations to be normal. In the following text we consider only models of AR type, therefore narrowing the model class. If the market is complete, in that there are so many generating sources of implicit uncertainty (the vector of innovations is of the same size as the vector of data on the left-hand size of a vector version of (1)) as there are data channels, this class will still contain the vector or many asset extension of the Brownian motion model. In a general vector form such stochastic process can be described by

$$\mathbf{\Phi}_{t+1} = \mathbf{A}\mathbf{\Phi}_t + \mathbf{\Sigma}\mathbf{e}_{t+1} \tag{2}$$

where $\mathbf{\Phi}_t$ is a general random vector (it has to be finite dimensional for the process to be Markovian) belonging to certain market data channels, containing among them the price Y_t, but generally also other data channels, possibly belonging to multiple assets, which is a case not further covered in this article, $\mathbf{A}, \mathbf{\Sigma}$ are matrices of unknown parameters and $\mathbf{e}_t$ is the innovation. Vectors and matrices are denoted by a bold font. The model is of arbitrary finite order in time, since in the data vector $\mathbf{\Phi}_{t-1}$ values with time lags greater than one are included by extending the model as shown in [13]. Because the matrix $\mathbf{\Sigma}$ is a square root of a positive definite matrix, we can reparametrize the model as

$$\mathbf{\Phi}_t = [\mathbf{I} - \mathbf{B}]\,\mathbf{\Phi}_t + \mathbf{B}\mathbf{A}\mathbf{\Phi}_{t-1} + \mathbf{D}^{\frac{1}{2}}\mathbf{e}_t \tag{3}$$

where $\mathbf{B}$ is a lower triangular matrix with units on the diagonal and $\mathbf{D}^{\frac{1}{2}}$ is a diagonal matrix. Thus we have effectively decoupled the vector process into individual AR processes with scalar left-hand side, so that without a loss of generality we can work with a model, where the innovation is a single random variable.

1.3 Proposing a model with Laplace innovations

In a well established practice the innovation is taken to be normally distributed, usually because of computational tractability of estimation of the parameters. Such a model for financial data in a Bayesian setup has been proposed and tested, see [13]. Koenker and Bassett show in their nowadays classic article [14] the dangers of such assumption for a linear regression model. Although correct modeling, if there is such, can resolve the heavy tailed behavior of market price difference distribution a "physical" explanation of the process of price creation would be needed to justify the model used. If we are not certain about the model, a more robust version may be more adequate. The efficiency of different estimators of parameters of linear regression model can be looked up in the above mentioned article [14] in Table I. These estimators are usually used by the robust statistics community in order to produce robust point estimates of the parameters in the model. Bayesian statistics on the other hand allow us to have access to the entire posterior likelihood of the parameters conditioned on the data, if we choose the distribution of the innovations and choose a Bayesian conjugate or improper prior. As shown in [14] median is a maximum a posteriori likelihood estimate in a linear regression model with Laplace errors. Inspired by the result I propose in the text the use of Laplace distribution in the AR model.

Distributions with heavier tails could be used to obtain an even more robust model, but as we shall see, such choice would make the Bayesian estimation too difficult and we see from [14] that the increase in efficiency of the estimation may be quite large even when real innovations have very heavy tails.

Example 1. For illustration of the usefulness of such an estimation at the cost of higher computational burden, we now look at maximum likelihood estimation of the mentioned model, where improper prior for the parameters was chosen. I have generated a sample price of 5000 samples from a model

$$Y_{t+1} = 0.97Y_t + e_{t+1} \tag{4}$$

with a starting condition $y_0 = 0$, where e_{t+1} is Student distributed with $\nu = 4, \mathrm{var}\,[e_{t+1}] = 1, \mu = 0$ with the parameters not chosen at random, but according to a distribution fit to price differences presented in [9]. Afterwards an estimation of the single parameter α was performed in a model

$$Y_{t+1} = \alpha Y_t + e_{t+1} \tag{5}$$

where the innovation was first taken to have normal and then Laplace distribution. In the first case the maximum likelihood estimate is of the form

$$\hat{\alpha}_{GML} = \frac{\sum_{t=2}^{T} y_t y_{t-1}}{\sum_{t=2}^{T} y_{t-1}^2} \tag{6}$$

where GML stands for Gaussian maximum likelihood, LML will be Laplace maximum likelihood and y_t are the realizations of random variables Y_t for $t \in \{1, \ldots, T\}$. In the second case the estimate comes down to computing

$$\hat{\alpha}_{LML} = \arg\min_{\alpha} \left[\sum_{t=2}^{T} |y_t - \alpha y_{t-1}| \right] \tag{7}$$

where the minimum is attained at one of the points, where the individual absolute value terms in the sum are equal to zero – $\alpha_t = \frac{y_t}{y_{t-1}}$ $t \in \{1, \ldots, T\}$. For such estimates I've evaluated a Monte Carlo variance test, where estimates GML and LML were computed for each of the m time-series of length n and empirical median, mean and variance were computed for these m samples. The results are summarized in the following table

	Gaussian			Laplacian		
(m,n)	median	mean	variance	median	mean	variance
$(4000, 50)$	0.9555	0.9399	$3.3 \cdot 10^{-3}$	0.9601	0.9473	$3.0 \cdot 10^{-3}$
$(2000, 100)$	0.9639	0.9550	$11 \cdot 10^{-4}$	0.9651	0.9594	$9.91 \cdot 10^{-4}$
$(1000, 200)$	0.9665	0.9615	$4.68 \cdot 10^{-4}$	0.9676	0.9639	$4.26 \cdot 10^{-4}$
$(1000, 500)$	0.9690	0.9668	$1.38 \cdot 10^{-4}$	0.9689	0.9677	$1.26 \cdot 10^{-4}$

Table 1: Empirical variance of estimates of parameter α considering Gaussian and Laplace noises.

2 Modeling

2.1 Bayesian basics

I will now focus on the Bayesian estimation of parameters in a model, where the parameters are constant in time. In such a model, when new data arrives the parameter distribution is updated according to the Bayes rule

$$f(\boldsymbol{\theta}|\mathbf{d}_t, \mathcal{F}_{t-1}) = \frac{f(\mathbf{d}_t|\boldsymbol{\theta}, \mathcal{F}_{t-1}) f(\boldsymbol{\theta}|\mathcal{F}_{t-1})}{\int_\Omega f(\mathbf{d}_t|\boldsymbol{\theta}, \mathcal{F}_{t-1}) f(\boldsymbol{\theta}|\mathcal{F}_{t-1}) d\boldsymbol{\theta}} \tag{8}$$

where Ω is the parameter space of parameters $\boldsymbol{\theta}$, $\mathbf{d}_t = (y_t, \phi_{t;1}, \phi_{t;2}, \ldots)$ is the newly arrived data vector realization and $\mathcal{F}_{t-1}$ is a filtration containing all the older information – in case of a data driven model it is fully determined by previous realizations of data. $f(\cdot)$ are probability density functions differentiated by their argument (including condition).

An update in a model with innovations of Gaussian form has been described in [15]. In such a case for the model to be analytically tractable, we choose the prior distribution of the parameters to be of the Gauss-Inverse-Wishart (Normal-Inverse-Gamma) form. Such a choice reduces the data update step to a simple algebraic update of sufficient statistics $\mathbf{V}$ and ν appearing in the GiW probability density. Such a reduction is caused by two important properties of the GiW probability density function:

- The density is a product of an exponential function with factors independent of the data, therefore the data update reduces to addition of the exponents of the two factors on the right-hand side of (8).
- The quadratic form in the exponents of both these factors allows for collapsing all the data entries into sufficient statistic $\mathbf{V}$.

In case of a model with Laplace innovations, the first property of the previous list is preserved, while the second one doesn't hold anymore as we shall see. At time t we can rewrite the model from the AR form

$$Y_t = \boldsymbol{\alpha}' \boldsymbol{\Phi}_t + \sigma e_t \tag{9}$$

where $'$ means transposition and the regression vector $\boldsymbol{\Phi}_t$ has length k, to a density form as

$$f(\boldsymbol{y_t}|\boldsymbol{\alpha}, \sigma, \boldsymbol{\phi}_t, \mathcal{F}_{t-p-1}) = \frac{1}{2\sigma} \exp\left[-\frac{1}{\sigma} |y_t - \boldsymbol{\alpha}' \boldsymbol{\phi}_t| \right] \tag{10}$$

so that this is the first factor in the numerator on the right-hand side of (8) and p is the model order. The second factor needs to be chosen in a form Bayesian conjugate to the model just mentioned, which means it has to stay form invariant after the Bayesian estimation step. Such a distribution has to be of the form

$$f(\boldsymbol{\alpha},\sigma|\mathcal{F}_0)=\frac{1}{I\sigma^{\nu}}\exp\left[-\frac{1}{\sigma}\sum_i|r_i-\mathbf{s}_i'\boldsymbol{\alpha}|\right] \tag{11}$$

where I is the normalization factor, $(r_i,\mathbf{s}_i)\in\mathbb{R}^{k+1}$ and i is from an arbitrary finite index set. As we proceed with the data update, we see numbers $r_i,\mathbf{s}_i$ are in the role of data realizations of the modelled variable so that after the first data update by (10) at time $t=p$ we obtain the posterior distribution of the parameters conditioned on data

$$f(\boldsymbol{\alpha},\sigma|\mathcal{F}_1)=\frac{1}{I_p\sigma^{\nu+1}}\exp\left[-\frac{1}{\sigma}\left[\sum_i|r_i-\mathbf{s}_i'\boldsymbol{\alpha}|+|y_p-\boldsymbol{\phi}_p'\boldsymbol{\alpha}|\right]\right] \tag{12}$$

which after reindexing could be rewritten in the same form as (11). For the sake of simple notation I will from now on consider the prior to be improper, so that $a_i,\mathbf{b}_i$ will not appear in the formulas and the only entries will be data.

2.2 Computing the normalization factor – integration over parameter space

As stated above, the Bayesian estimation will allow us to construct an entire posterior distribution of the parameters. As new data arrive we have to perform the data update of the distribution and because the terms with absolute value in the exponent of the posterior distribution don't collapse into a low-dimensional sufficient statistic, we have to remember all the previous data obtained in a form of a table and we are practically done with computations except for the terms containing an integral over the parameter space – namely I_p in (12) and the denominator on the right-hand side of (8). We shall now study such an integral in greater detail. In a case of multivariate parameter $\boldsymbol{\alpha}$, the integral can be written as a sum of terms without the absolute value in exponent

$$\begin{aligned}\int_\Omega\frac{1}{I_{\tau-1}\sigma^{\tau}}\exp\left[-\frac{1}{\sigma}\sum_{t=2}^{\tau}\left|y_t-\boldsymbol{\phi}_t'\boldsymbol{\alpha}\right|\right]d\sigma d\boldsymbol{\alpha}=&\int_0^\infty\left(\int_{R_1}\frac{1}{I_{\tau-1}\sigma^{\tau}}\exp\left[-\frac{1}{\sigma}\left(\bar{y}_{\tau;R_1}-\bar{\boldsymbol{\phi}}_{\tau;R_1}'\boldsymbol{\alpha}\right)\right]d\boldsymbol{\alpha}\right)d\sigma+\\&+\int_0^\infty\left(\int_{R_2}\frac{1}{I_{\tau-1}\sigma^{\tau}}\exp\left[-\frac{1}{\sigma}\left(\bar{y}_{\tau;R_2}-\bar{\boldsymbol{\phi}}_{\tau;R_2}'\boldsymbol{\alpha}\right)\right]d\boldsymbol{\alpha}\right)d\sigma+\cdots\end{aligned} \tag{13}$$

where

$$\begin{aligned}R_i=\left\{\boldsymbol{\alpha}:\forall j\in\mathcal{J}_i,y_j-\boldsymbol{\phi}_j'\boldsymbol{\alpha}\geq 0;\forall l\in\mathcal{L}_i,y_l-\boldsymbol{\phi}_l'\boldsymbol{\alpha}<0\right\}\\ i\in\{1,2,\ldots,N\};\mathcal{J}_i,\mathcal{L}_i\subset\{1,\ldots,\tau\};\mathcal{L}_i=\mathcal{J}_i^C\end{aligned} \tag{14}$$

where ^{C} stands for complement in the set of indices $\{1,\ldots,\tau\}$ and

$$\bar{y}_{\tau;R_i}=\sum_{t=p}^{\tau}(-1)^{\mathbb{I}(t\in\mathcal{L}_i)}y_t\qquad\bar{\phi}_{\tau;R_i;j}=\sum_{t=p}^{\tau}(-1)^{\mathbb{I}(t\in\mathcal{L}_i)}\phi_{t;j} \tag{15}$$

The parameter space is therefore divided by hyperplanes $y_t-\boldsymbol{\phi}_t'\boldsymbol{\alpha}=0$ into convex polytopes R_i, which then serve as a domain of integration for the individual integrals. Alternatively to the set of inequalites, the polytope can be represented by all its vertices $\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_k$ or a complex combination of a vertex and vectors leading from one vertex to the next $\mathbf{u}_1=\mathbf{v}_2-\mathbf{v}_1,\mathbf{u}_2=\mathbf{v}_3-\mathbf{v}_1,\ldots,\mathbf{u}_{k-1}=\mathbf{v}_k-\mathbf{v}_1,$. The vertices could be obtained by solving the individual sets of linear inequalities, but such a serial computation is not at all effective, so we will state a more effective algorithm for establishing the region later. Now suppose, we know all the vertices of a given polytope and have to solve one of the integrals appearing on the right-hand side of (13)

$$\int_{\substack{\mathbf{v}_1+x_1\mathbf{u}_1+x_2\mathbf{u}_2+\ldots+x_k\mathbf{u}_k\\ \forall i\ x_i\geq 0,\sum_{i=1}^k x_i\leq 1,\sigma\in(0,\infty)}}\frac{1}{I_{\tau-1}\sigma^{\tau}}\exp\left[-\frac{1}{\sigma}\left(\bar{y}_t-\bar{\boldsymbol{\phi}}_t'\boldsymbol{\alpha}\right)\right]d\sigma d\alpha=\spadesuit \tag{16}$$

The integral over $\boldsymbol{\alpha}$ can be solved by the method presented in [16], but such integration only works for n-polytopes – n dimensional polytopes with $n+1$ vertices, because then the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}$ are linearly independent. If the an n-dimensional polytope has more than $n+1$ vertices, we must first find its triangulation. A method of triangulation is described in [17]. Once we have a triangulation into n-polytopes we can use the substitution theorem to transform the integral into an integral over an n-simplex coordinates $x_1, \ldots, x_k$, by a linear transform

$$\alpha_i(\mathbf{x}) = u_{i0} + x_1 u_{i1} + x_2 u_{i2} + \ldots + x_k u_{ik} \tag{17}$$

where

$$u_{i0} = v_{i1} \qquad u_{ij} = v_{ij} - v_{i0} \qquad i,j \in \{1,\ldots,k\} \tag{18}$$

we obtain

$$\spadesuit = \int_0^\infty \frac{1}{I_{\tau-1}\sigma^\tau} \exp \underbrace{\left[-\frac{\bar{y}_t + \sum_{j=1}^k \bar{\phi}_{t;j} u_{j0}}{\sigma}\right]}_{-a_0/\sigma} \int_0^1 \int_0^{\Phi_1} \int_0^{\Phi_2} \cdots \int_0^{\Phi_{k-1}} \prod_{i=1}^k \exp \underbrace{\left[\frac{\sum_{j=1}^k \bar{\phi}_{t;j} u_{ji}}{\sigma}\right]}_{a_i/\sigma} x_i \, |J| \, d\mathbf{x} d\sigma \tag{19}$$

where the Jacobian

$$|J| = \begin{vmatrix} u_{11} & u_{12} & \cdots & u_{1k} \\ \vdots & \vdots & & \vdots \\ u_{k1} & u_{k2} & \cdots & u_{kk} \end{vmatrix} \tag{20}$$

and

$$\Phi_i = 1 - x_1 - x_2 - \ldots - x_i \qquad i \in \{1,2,\ldots,k-1\} \tag{21}$$

Such an integral can already be analytically solved and the solution can be proved to be

$$\spadesuit = \frac{\Gamma(\tau-k-1)\,|J|}{I_{\tau-1}} \sum_{i=1}^k \left[\frac{1}{a_i \, (a_i - a_0)^{\tau-k-1}} \prod_{\substack{j=1,\ldots,k \\ i \neq j}} \frac{1}{(a_i - a_j)} \right] \tag{22}$$

where the a_is have been defined in equation (19).

2.3 Efficient algorithm, computational issues and future work

For being able to integrate over the entire parameter space, we have to be able to efficiently split the space into a complex of convex polytopes. Such a split could be obtained by serial solution of the sets of inequalities in equation (13), but such a procedure would be very inefficient. Instead, we can proceed with a splitting algorithm presented in [18]. We first need a Hasse diagram representation of the complex – for each convex polyhedron a collection of sets of lower d-dimensional polytopes N_d, where $d \in \{0,1,\ldots,k\}$ and arrows marking subordinations of polytopes between neighboring sets N_{d-1}, N_d. From such a representation we can proceed with the algorithm. For each polytope in the complex we do

begin

1. Classify the position of all the vertices relative to the splitting hyperplane and mark them $\boxed{+}$ and $\boxed{-}$. Set $i = 1$.
2. For each $p \in N_i$ do:
   - If none of the facets ($f \in N_{i-1}$ being connected by an arrow to p) is $\boxed{+}$($\boxed{-}$), set p to be $\boxed{-}$($\boxed{+}$) and goto 3.
   - Create new $i-1$ dimensional polytope q, set it as $\boxed{=}$ and connect it to all $i-2$ dimensional polytopes in p classified $\boxed{=}$. Create two polytopes $p+$ marked $\boxed{+}$ and $p-$ marked $\boxed{-}$, connect it to q and all the $i+1$ polytopes connected to p. Connect to $p+$($p-$) all the $i-1$ facets of p marked $\boxed{+}$($\boxed{-}$). Remove p from N_i.
3. If $i < k$ set $i = i+1$ and goto 2.
4. Compute the geometric coordinates of vertices marked $\boxed{=}$.

end

With the use of the algoritm come certain computational issues that I have not resolved yet. At the beginning we need a starting complex of convex polytopes to split. Such a complex can be obtained by solving the set of the first k non-degenerate hyperplane conditions and obtaining their intersection and

all the lines going through that intersection. As can be seen, the polytopes of such complex will not be bounded. We can resolve this by cutting these polytopes off at some point. Such a cutoff has to be carefully considered and its incorporation into the above algorithm is still not fully discovered.

Also when new conditions (data) arrive, some polytopes obtained by the splitting algorithm can be very little. A way of treating these regions is given in [18]. Another numerical issue can arrise when we integrate over very little region high function values at the far tails where we integrate little function values over large space regions. The influence of these numerical inefficiencies have to be considered as well.

Finally, it would be nice to incorporate the triangulation of k-dimensional polytopes which are not k-polytopes into the algorithm presented above to reduce computational needs. Again such an incorporation has been left for future study.

3 Acknowledgment

I would like to thank Ministry of Education, Youth and Sports of the Czech Republic for funding this research through project DAR 1M0572. I would also like to thank my supervisor Dr. Miroslav Kárný for his helpful comments on the article.

References

[1] L. Bachelier, *Théorie de la speculation.* Paris: Gauthier-Villars, 1900, reprinted in P.H.Cootner: The Random Character of Stock Market Prices(1964).

[2] R. Merton, "Theory of rational option pricing," *The Bell Journal of Economics and Management Science*, vol. 4, no. 1, pp. 141–183, 1973.

[3] F. Black and M. Scholes, "The pricing of options and corporate liabilities," *Journal of Political Economy*, vol. 81, no. 3, pp. 637–654, 1973.

[4] I. Karatzas and S. Shreve, *Methods of Mathematical Finance.* Springer, 1998.

[5] E. Fama, "The behavior of stock market prices," *Journal of Business*, vol. 38, p. 34105, 1965.

[6] P. Samuelson, "Proof that properly anticipated prices fluctuate randomly," *Industrial Management Review*, vol. 6, pp. 44–49, 1965.

[7] G. Box and G. Jenkins, *Time Series Analysis.* San Francisco, U.S.A.: Holden-Day, 1970.

[8] P. Billingsley, *Convergence of Probability Measures*, 2nd ed. Wiley, 1999.

[9] J. Bouchaud and M. Potters, *Theory of financial risk and derivative pricing: from statistical physics to risk management*, 2nd ed. Cambridge University Press, 2003.

[10] R. Cont and P. Tankov, *Financial Modelling With Jump Processes.* CRC Press, 2004.

[11] F. Canova and T. Ito, "The time-series properties of the risk premium in the yen/dollar exchange market," *Journal of Applied Econometrics*, vol. 6, pp. 125–142, 1991.

[12] N. Sarantis, "On the short-term predictability of exchange rates: A bvar time-varying parameters approach," *Journal of Banking & Finance*, vol. 30, pp. 2257–2279, 2006.

[13] J. Šindelář, "Study of bvar(p) process applied to u.s. commodity market data," in *World Academy of Science, Engineering and Technology*, C. Ardil, Ed., vol. 58, 2009, pp. 424–435.

[14] R. Koenker and G. Bassett, "Regression quantiles," *Econometrica*, vol. 46, no. 1, pp. 33–50, 1978.

[15] M. Kárný, J. Böhm, T. Guy, L. Jirsa, I. Nagy, P. Nedoma, and L. Tesař, *Optimized Bayesian Dynamic Advising, Theory and Algorithms.* Springer, 2005.

[16] H. Rathod and G. Rao, "Integration of polynomials over n-dimensional linear polyhedra," *Comuputers & Structures*, vol. 65, no. 6, pp. 829–847, 1997.

[17] J. Cohen and T. Hickey, "Two algorithms for detemining volumes of convex polyhedra," *Journal of the Association for Computing Machinery*, vol. 26, pp. 401–414, 1979.

[18] C. Bajaj and V. Pascucci, "Splitting a complex of convex polytopes in any dimension," in *Proceedings of the twelfth annual symposium on Computational geometry.* ACM New York, 1996, pp. 88–97.