# Advances in robust AR model research

Jan Šindelář

Department of Adaptive Systems, UTIA Academy of Sciences of the Czech Republic

Department of Probability and Mathematical Statistics, Charles University

e-mail: jenik.sindelar@gmail.com

Outline:

- Bayesian setup
- Modelling uncertainty in financial data with auto-regression models
- Robust version of AR model
- Current state of the solution

Outline:

- Bayesian setup
- Modelling uncertainty in financial data with auto-regression models
- Robust version of AR model
- Current state of the solution

Outline:

- Bayesian setup
- Modelling uncertainty in financial data with auto-regression models
- Robust version of AR model
- Current state of the solution

Outline:

- Bayesian setup
- Modelling uncertainty in financial data with auto-regression models
- Robust version of AR model
- Current state of the solution

We try to model a discrete random process $Y = (Y_0, Y_1, \ldots, Y_T)$.

Assumption: probability measure $\mu$ to be absolutely continuous with respect to underlying Lebesgue measure – density exists

$$\mu(dY) = f(Y)\lambda(dY) \tag{1}$$

Uncertainty in $Y$ can be described by observable environment model and unobservable parameters $\theta$. Considering a parameterized model for individual $Y_t$ the previously described density factorizes into

$$f(Y) = \underbrace{f(\theta)}_{prior} \prod_{t \in T^*} \underbrace{f(Y_t|\mathcal{F}_t, \theta)}_{model} \tag{2}$$

In such cases new data "describe" the properties of the parameters in a way given by Bayes formula

$$f(\theta|Y_t, \mathcal{F}_{t-1}) = \frac{f(Y_t|\theta, \mathcal{F}_{t-1}) f(\theta|\mathcal{F}_{t-1})}{\int_\Omega f(Y_t|\theta, \mathcal{F}_{t-1}) f(\theta|\mathcal{F}_{t-1}) d\theta} \tag{3}$$

where $\Omega$ is the parameter space.

We try to model a discrete random process $Y = (Y_0, Y_1, \dots, Y_T)$.

Assumption: probability measure $\mu$ to be absolutely continuous with respect to underlying Lebesgue measure – density exists

$$\mu(dY) = f(Y)\lambda(dY) \tag{1}$$

Uncertainty in $Y$ can be described by observable environment model and unobservable parameters $\theta$. Considering a parameterized model for individual $Y_t$ the previously described density factorizes into

$$f(Y) = \underbrace{f(\theta)}_{prior} \prod_{t \in T^*} \underbrace{f(Y_t | \mathcal{F}_t, \theta)}_{model} \tag{2}$$

In such cases new data "describe" the properties of the parameters in a way given by Bayes formula

$$f(\theta | Y_t, \mathcal{F}_{t-1}) = \frac{f(Y_t | \theta, \mathcal{F}_{t-1}) f(\theta | \mathcal{F}_{t-1})}{\int_\Omega f(Y_t | \theta, \mathcal{F}_{t-1}) f(\theta | \mathcal{F}_{t-1}) d\theta} \tag{3}$$

where $\Omega$ is the parameter space.

Ideally the model is given by the physics of the system – for financial data Econophysics. If it is not, we have to choose rich and still computationally feasible model.

Possible and often used is linear auto-regression model of the random process $Y_t$

$$\underbrace{\begin{bmatrix} \mathbf{D}_{t+1} \\ \mathbf{D}_t \\ \mathbf{D}_{t-1} \\ \vdots \\ \mathbf{D}_{t+1-p} \\ 1 \end{bmatrix}}_{Y_{t+1}} = \underbrace{\begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 & & \mathbf{A}_{p-1} & \mathbf{A}_p & \mathbf{c} \\ \mathbf{I} & \mathbf{0} & & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \cdots & \vdots & & \\ \mathbf{0} & \mathbf{0} & & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & & \mathbf{0} & \mathbf{0} & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \mathbf{D}_t \\ \mathbf{D}_{t-1} \\ \mathbf{D}_{t-2} \\ \vdots \\ \mathbf{D}_{t-p} \\ 1 \end{bmatrix}}_{Y_t} + \underbrace{\begin{bmatrix} \boldsymbol{\Sigma} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \cdots & \mathbf{0} \\ \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} \mathbf{e}_{t+1} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}}_{e_{t+1}}$$

So rewritten into the original notation, we have

$$Y_{t+1} = AY_t + \Sigma e_{t+1} \tag{4}$$

where $\theta = A, \Sigma$ are the parameters $e_t$ is the previously mentioned observable, modelled uncertainty or innovation.

Ideally the model is given by the physics of the system – for financial data Econophysics. If it is not, we have to choose rich and still computationally feasible model.

Possible and often used is linear auto-regression model of the random process $Y_t$

$$\underbrace{\begin{bmatrix} \mathbf{D}_{t+1} \\ \mathbf{D}_t \\ \mathbf{D}_{t-1} \\ \vdots \\ \mathbf{D}_{t+1-p} \\ 1 \end{bmatrix}}_{Y_{t+1}} = \underbrace{\begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 & & \mathbf{A}_{p-1} & \mathbf{A}_p & \mathbf{c} \\ \mathbf{I} & \mathbf{0} & & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \cdots & \vdots & & \\ \mathbf{0} & \mathbf{0} & & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & & \mathbf{0} & \mathbf{0} & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \mathbf{D}_t \\ \mathbf{D}_{t-1} \\ \mathbf{D}_{t-2} \\ \vdots \\ \mathbf{D}_{t-p} \\ 1 \end{bmatrix}}_{Y_t} + \underbrace{\begin{bmatrix} \mathbf{\Sigma} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \cdots & \mathbf{0} \\ \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} \mathbf{e}_{t+1} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}}_{e_{t+1}}$$

So rewritten into the original notation, we have

$$Y_{t+1} = AY_t + \Sigma e_{t+1} \tag{4}$$

where $\theta = A, \Sigma$ are the parameters $e_t$ is the previously mentioned observable, modelled uncertainty or innovation.

In a standard setup studied at our department for a long time already $e_t \sim \mathcal{N}(0, I)$ and Bayesian update becomes simple algebraic operation on sufficient statistic, if self-reproducing prior is chosen.

I've been playing with the standard setup for quite a long time: forgetting, multi-variate auto-regression, order and structure selection, log-normal instead of normal prices, multi-step ahead predictions.

Performed experiments on real exchange data of daily commodity prices to compare an AR model with models used in todays financial mathematics.

In a standard setup studied at our department for a long time already $e_t \sim \mathcal{N}(0, I)$ and Bayesian update becomes simple algebraic operation on sufficient statistic, if self-reproducing prior is chosen.

I've been playing with the standard setup for quite a long time: forgetting, multi-variate auto-regression, order and structure selection, log-normal instead of normal prices, multi-step ahead predictions.

Performed experiments on real exchange data of daily commodity prices to compare an AR model with models used in todays financial mathematics.

In a standard setup studied at our department for a long time already $e_t \sim \mathcal{N}(0, I)$ and Bayesian update becomes simple algebraic operation on sufficient statistic, if self-reproducing prior is chosen.

I've been playing with the standard setup for quite a long time: forgetting, multi-variate auto-regression, order and structure selection, log-normal instead of normal prices, multi-step ahead predictions.

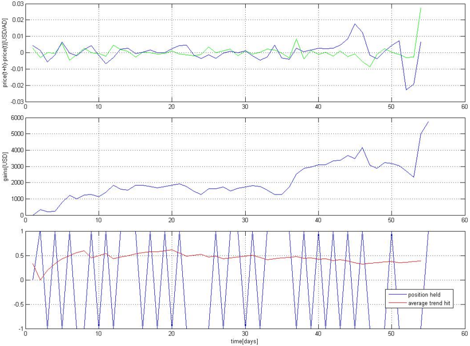
Performed experiments on real exchange data of daily commodity prices to compare an AR model with models used in todays financial mathematics.

In a standard setup studied at our department for a long time already $e_t \sim \mathcal{N}(0, I)$ and Bayesian update becomes simple algebraic operation on sufficient statistic, if self-reproducing prior is chosen.

I've been playing with the standard setup for quite a long time: forgetting, multi-variate auto-regression, order and structure selection, log-normal instead of normal prices, multi-step ahead predictions.

Performed experiments on real exchange data of daily commodity prices to compare an AR model with models used in todays financial mathematics.

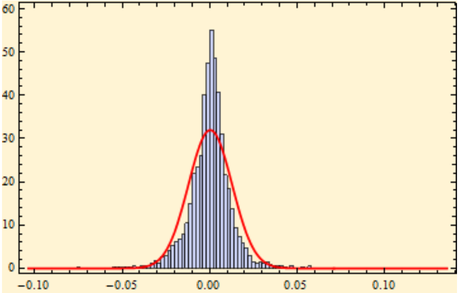
We experimented with simulated trading:



Motivation: Log-returns

$$LR_i = \ln y_{t+i+1} - \ln y_{t+i} \tag{5}$$

of prices are leptokurtically distributed.




Best fit ≈ Student with $\nu = 4$ (Bouchaud).

If we use any regression type model, estimates we get are heavily influenced by outliers/non-normality (Koneker& Bassett)

TABLE I
EMPIRICAL VARIANCES OF SOME ALTERNATIVE LOCATION ESTIMATORS[a]
(Sample Size 20)

| Estimators | Distributions: Normal | 10% $3\sigma$[b] | 10% $10\sigma$[c] | Laplace | Cauchy |
|---|---|---|---|---|---|
| Mean | 1.00 | 1.88 | 11.54 | 2.10 | 12,548.0 |
| 10% trimmed mean | 1.06 | 1.31 | 1.46 | 1.60 | 7.3 |
| 25% trimmed mean | 1.20 | 1.41 | 1.47 | 1.33 | 3.1 |
| Median | 1.50 | 1.70 | 1.80 | 1.37 | 2.9 |
| Gastwirth[d] | 1.23 | 1.45 | 1.51 | 1.35 | 3.1 |
| Trimean[e] | 1.15 | 1.37 | 1.48 | 1.43 | 3.9 |

[a] Abstracted from Exhibit 5 in Andrews, *et al.* [3].
[b] Gaussian Mixture: $.9\Phi(1)+.1\Phi(3)$.
[c] Gaussian Mixture: $.9\Phi(1)+.1\Phi(10)$.
[d] $\tilde{\beta}=.3\beta^*(1/3)+.4\beta^*(1/2)+.3\beta^*(2/3)$, where $\beta^*(\theta)$ is the $\theta$th sample quantile.
[e] $\tilde{\beta}=1/4\beta^*(1/4)+1/2\beta^*(1/2)+1/4\beta^*(3/4)$.

It might be useful to replace normal with other – I will propose Laplace (double exponential).

Motivation: In linear regression model median is a maximum likelihood estimate if innovations are Laplace, mean if Gaussian.

In a model with constant parameters, we use Bayesian data update

$$f(\theta|Y_t, \mathcal{F}_{t-1}) = \frac{f(Y_t|\theta, \mathcal{F}_{t-1})f(\theta|\mathcal{F}_{t-1})}{\int_{\Omega} f(Y_t|\theta, \mathcal{F}_{t-1})f(\theta|\mathcal{F}_{t-1})d\theta} \tag{6}$$

In a model with Gaussian innovations and GiW (NiG) prior the estimation has two important properties

- Exponential form of density – transforms multiplication into summation of exponents
- Quadratic (polynomial) form in the exponent conserves form when summed with other quadratic form (polynomial of same order)

in a model with Laplace innovations and proper prior the first property holds, while the second one fails

In a model with constant parameters, we use Bayesian data update

$$f(\theta|Y_t, \mathcal{F}_{t-1}) = \frac{f(Y_t|\theta, \mathcal{F}_{t-1})f(\theta|\mathcal{F}_{t-1})}{\int_\Omega f(Y_t|\theta, \mathcal{F}_{t-1})f(\theta|\mathcal{F}_{t-1})d\theta} \quad (6)$$

In a model with Gaussian innovations and GiW (NiG) prior the estimation has two important properties

- Exponential form of density – transforms multiplication into summation of exponents
- Quadratic (polynomial) form in the exponent conserves form when summed with other quadratic form (polynomial of same order)

in a model with Laplace innovations and proper prior the first property holds, while the second one fails

In a model with constant parameters, we use Bayesian data update

$$f(\theta|Y_t, \mathcal{F}_{t-1}) = \frac{f(Y_t|\theta, \mathcal{F}_{t-1})f(\theta|\mathcal{F}_{t-1})}{\int_{\Omega} f(Y_t|\theta, \mathcal{F}_{t-1})f(\theta|\mathcal{F}_{t-1})d\theta} \tag{6}$$

In a model with Gaussian innovations and GiW (NiG) prior the estimation has two important properties

- Exponential form of density – transforms multiplication into summation of exponents
- Quadratic (polynomial) form in the exponent conserves form when summed with other quadratic form (polynomial of same order)

in a model with Laplace innovations and proper prior the first property holds, while the second one fails

In case of AR model, where $e_t$ is Laplace(0,1) white noise

$$Y_t = \boldsymbol{\alpha}'\boldsymbol{\Phi}_t + \sigma e_t \tag{7}$$

the model density is

$$f(\boldsymbol{y_t}|\boldsymbol{\alpha}, \sigma, \phi_t, \mathcal{F}_{t-p-1}) = \frac{1}{2\sigma} \exp\left[-\frac{1}{\sigma}\left|y_t - \boldsymbol{\alpha}'\phi_t\right|\right] \tag{8}$$

and Bayesian self-reproducing prior

$$f(\alpha, \sigma|\mathcal{F}_0) = \frac{1}{I\sigma^{\nu}} \exp\left[-\frac{1}{\sigma}\sum_{i=1}^{\nu}\left|r_i - \mathbf{s}_i'\alpha\right|\right] \tag{9}$$

where $(r_i, \mathbf{s}_i) \in \mathbb{R}^{k+1}$. After first data update

$$f(\alpha, \sigma|\mathcal{F}_p) = \frac{1}{I_p\sigma^{\nu+1}} \exp\left[-\frac{1}{\sigma}\left[\sum_{i=1}^{\nu}\left|r_i - \mathbf{s}_i'\alpha\right| + \left|y_p - \phi_p'\alpha\right|\right]\right] \tag{10}$$

which is the posterior. For simplicity from now on I consider improper prior, $r_i$, $\mathbf{s}_i$ disappear. As no low dimensional sufficient statistics appear, the only thing left is $I_p$. Good for structure estimation and further analytical computations (moments, etc.)

In case of AR model, where $e_t$ is Laplace(0,1) white noise

$$Y_t = \boldsymbol{\alpha}'\boldsymbol{\Phi}_t + \sigma e_t \tag{7}$$

the model density is

$$f(\boldsymbol{y_t}|\boldsymbol{\alpha}, \sigma, \phi_t, \mathcal{F}_{t-p-1}) = \frac{1}{2\sigma} \exp\left[-\frac{1}{\sigma}\left|y_t - \boldsymbol{\alpha}'\phi_t\right|\right] \tag{8}$$

and Bayesian self-reproducing prior

$$f(\boldsymbol{\alpha}, \sigma|\mathcal{F}_0) = \frac{1}{I\sigma^{\nu}} \exp\left[-\frac{1}{\sigma}\sum_{i=1}^{\nu}\left|r_i - \mathbf{s}_i'\boldsymbol{\alpha}\right|\right] \tag{9}$$

where $(r_i, \mathbf{s}_i) \in \mathbb{R}^{k+1}$. After first data update

$$f(\alpha, \sigma|\mathcal{F}_p) = \frac{1}{I_p\sigma^{\nu+1}} \exp\left[-\frac{1}{\sigma}\left[\sum_{i=1}^{\nu}\left|r_i - \mathbf{s}_i'\alpha\right| + \left|y_p - \phi_p'\alpha\right|\right]\right] \tag{10}$$

which is the posterior. For simplicity from now on I consider improper prior, $r_i$, $\mathbf{s}_i$ disappear. As no low dimensional sufficient statistics appear, the only thing left is $I_p$. Good for structure estimation and further analytical computations (moments, etc.)

In case of AR model, where $e_t$ is Laplace(0,1) white noise

$$Y_t = \boldsymbol{\alpha}' \boldsymbol{\Phi}_t + \sigma e_t \tag{7}$$

the model density is

$$f(\boldsymbol{y_t} | \boldsymbol{\alpha}, \sigma, \phi_t, \mathcal{F}_{t-p-1}) = \frac{1}{2\sigma} \exp\left[ -\frac{1}{\sigma} \left| y_t - \boldsymbol{\alpha}' \phi_t \right| \right] \tag{8}$$

and Bayesian self-reproducing prior

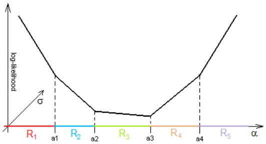
$$f(\boldsymbol{\alpha}, \sigma | \mathcal{F}_0) = \frac{1}{I\sigma^{\nu}} \exp\left[ -\frac{1}{\sigma} \sum_{i=1}^{\nu} \left| r_i - \mathbf{s}_i' \boldsymbol{\alpha} \right| \right] \tag{9}$$

where $(r_i, \mathbf{s}_i) \in \mathbb{R}^{k+1}$. After first data update

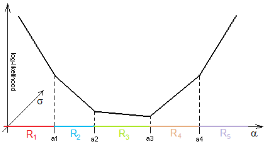
$$f(\boldsymbol{\alpha}, \sigma | \mathcal{F}_p) = \frac{1}{I_p \sigma^{\nu+1}} \exp\left[ -\frac{1}{\sigma} \left[ \sum_{i=1}^{\nu} \left| r_i - \mathbf{s}_i' \boldsymbol{\alpha} \right| + \left| y_p - \phi_p' \boldsymbol{\alpha} \right| \right] \right] \tag{10}$$

which is the posterior. For simplicity from now on I consider improper prior, $r_i, \mathbf{s}_i$ disappear. As no low dimensional sufficient statistics appear, the only thing left is $I_p$. Good for structure estimation and further analytical computations (moments, etc.)

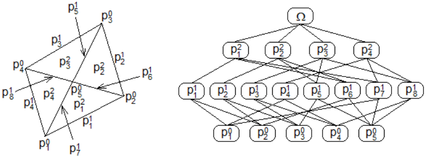
Graphically the posterior is an exponential function with exponent of the following type (here for one location parameter)



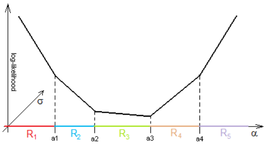
Graphically the posterior is an exponential function with exponent of the following type (here for one location parameter)



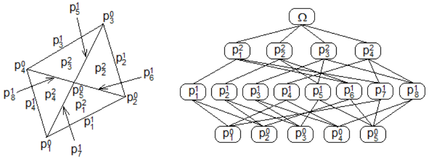

Algorithm for cutting parameter space(inspired by Bajaj, Pascucci): Based on a Hasse diagram representation



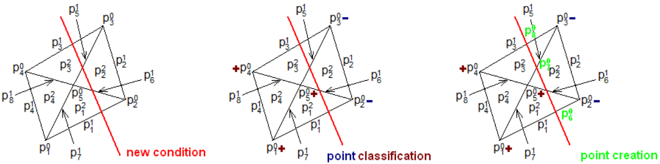
Graphically the posterior is an exponential function with exponent of the following type (here for one location parameter)



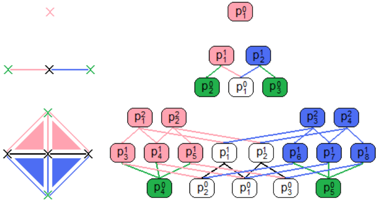

Algorithm for cutting parameter space(inspired by Bajaj, Pascucci): Based on a Hasse diagram representation




Algorithm procedure:

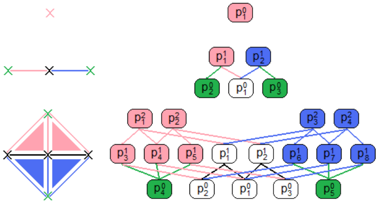


The algorithm is from Hasse to Hasse, so we need a Hasse diagram to start. We create the first iteration by representing the parameter space by a Hasse diagram



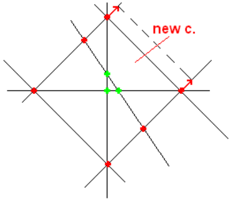
can be thought of as a prior knowledge about the parameters from a Bayesian viewpoint.

The algorithm is from Hasse to Hasse, so we need a Hasse diagram to start. We create the first iteration by representing the parameter space by a Hasse diagram

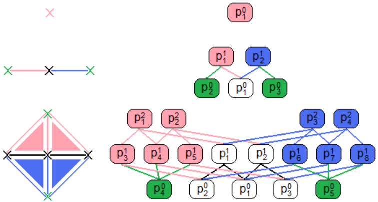



can be thought of as a prior knowledge about the parameters from a Bayesian viewpoint.

This is not necessary the best tractable solution. Original idea was to cut the space as mentioned, but introduce the idea of double points.

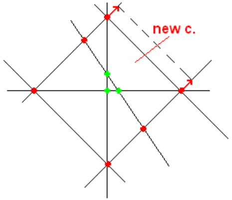


The algorithm is from Hasse to Hasse, so we need a Hasse diagram to start. We create the first iteration by representing the parameter space by a Hasse diagram



can be thought of as a prior knowledge about the parameters from a Bayesian viewpoint.

This is not necessary the best tractable solution. Original idea was to cut the space as mentioned, but introduce the idea of double points.




**Is there no obstacle?**

If we try to obtain $I_p$ in (8), we have to integrate the exponential over parameter space. We obtain an integral of the following type

$$I_p \propto \int_0^\infty e^{-a_0} \left( \int_0^1 e^{a_1 x_1} \left( \int_0^{1-x_1} e^{a_2 x_2} \cdots \left( \int_0^{1-x_1-x_2\ldots-x_n} e^{a_n x_n} dx_n \right) \cdots dx_2 \right) dx_1 \right) d\sigma \tag{11}$$

If we try to obtain $I_p$ in (8), we have to integrate the exponential over parameter space. We obtain an integral of the following type

$$I_p \propto \int_0^{\infty} e^{-a_0} \left( \int_0^1 e^{a_1 x_1} \left( \int_0^{1-x_1} e^{a_2 x_2} \cdots \left( \int_0^{1-x_1-x_2 \ldots -x_n} e^{a_n x_n} dx_n \right) \cdots dx_2 \right) dx_1 \right) d\sigma \tag{12}$$

Doesn't seem good, maybe $2^n$ terms, because if we integrate the first integral, we see that

$$\int_0^A e^{bx} dx = \frac{1}{b} \left[ e^{bA} - 1 \right] \tag{13}$$

If we try to obtain $I_p$ in (8), we have to integrate the exponential over parameter space. We obtain an integral of the following type

$$I_p \propto \int_0^\infty e^{-a_0} \left( \int_0^1 e^{a_1 x_1} \left( \int_0^{1-x_1} e^{a_2 x_2} \cdots \left( \int_0^{1-x_1-x_2\ldots-x_n} e^{a_n x_n} dx_n \right) \cdots dx_2 \right) dx_1 \right) d\sigma \tag{14}$$

Doesn't seem good, maybe $2^n$ terms, because if we integrate the first integral, we see that

$$\int_0^A e^{bx} dx = \frac{1}{b} \left[ e^{bA} - 1 \right] \tag{15}$$

The integrals can be computed, if we know the triangulation of the polyhedron. The result can be shown to be

$$\frac{\Gamma(\tau - k - 1) \, |J|}{I_{\tau-1}} \sum_{i=1}^{k} \left[ \frac{1}{a_i \, (a_i - a_0)^{\tau-k-1}} \prod_{\substack{j=1,\ldots,k \\ i \neq j}} \frac{1}{(a_i - a_j)} \right] \tag{16}$$

where the $a$'s are computed with the use of vertex coordinates and summed conditions.

If we try to obtain $I_p$ in (8), we have to integrate the exponential over parameter space. We obtain an integral of the following type

$$I_p \propto \int_0^{\infty} e^{-a_0} \left( \int_0^1 e^{a_1 x_1} \left( \int_0^{1-x_1} e^{a_2 x_2} \cdots \left( \int_0^{1-x_1-x_2\ldots-x_n} e^{a_n x_n} dx_n \right) \cdots dx_2 \right) dx_1 \right) d\sigma \tag{17}$$

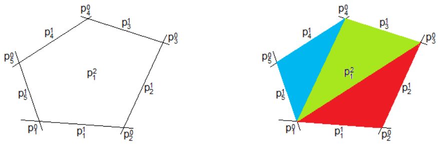
Doesn't seem good, maybe $2^n$ terms, because if we integrate the first integral, we see that

$$\int_0^A e^{bx} dx = \frac{1}{b}\left[ e^{bA} - 1 \right] \tag{18}$$

The integrals can be computed, if we know the triangulation of the polyhedron. The result can be shown to be

$$\frac{\Gamma(\tau - k - 1)\,|J|}{I_{\tau-1}} \sum_{i=1}^{k} \left[ \frac{1}{a_i\,(a_i - a_0)^{\tau-k-1}} \prod_{\substack{j=1,\ldots,k \\ i \neq j}} \frac{1}{(a_i - a_j)} \right] \tag{19}$$

where the $a$'s are computed with the use of vertex coordinates and summed conditions.
For triangulation:



I have used the theorem of Cohen & Hickey.

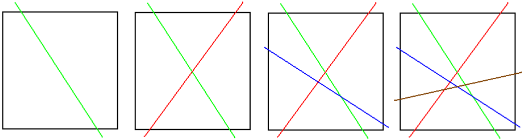
Bad news:
If $a_i = a_j$ for some $i \neq j$, the integral seems to diverge. It is a natural phenomenon.
Limit integral switching problem. Can be resolved using Taylor expansion.
For example, when $\alpha$ is 4-D, when $a_1 = a_2 = a_3$

$$\left[\frac{1}{2} - \frac{1}{a_3} - \frac{1}{a_3 - a_4} + \frac{1}{a_3^2} + \frac{1}{(a_3 - a_4)^2} + \frac{1}{a_3(a_3 - a_4)}\right] \frac{\Gamma(\tau - 4)|J|}{a_3(a_3 - a_4)(a_0 - a_3)^{\tau - 4} I_{\tau - 1}} \tag{20}$$

Bad news:
If $a_i = a_j$ for some $i \neq j$, the integral seems to diverge. It is a natural phenomenon.
Limit integral switching problem. Can be resolved using Taylor expansion.
For example, when $\alpha$ is 4-D, when $a_1 = a_2 = a_3$

$$\left[\frac{1}{2} - \frac{1}{a_3} - \frac{1}{a_3 - a_4} + \frac{1}{a_3^2} + \frac{1}{(a_3 - a_4)^2} + \frac{1}{a_3(a_3 - a_4)}\right] \frac{\Gamma(\tau - 4)|J|}{a_3(a_3 - a_4)(a_0 - a_3)^{\tau - 4} I_{\tau - 1}} \quad (21)$$

Dividing the space:



1 parameters: $1 \to 2 \to 3 \to 4 \to 5$
2 parameters: $1 \to 2 \to 4 \to 7 \to 11$
3 parameters: $1 \to 2 \to 4 \to 8 \to 15$
etc.
The sum $S_n$ of the series can be computed. For large $n$ it is proportional to $n^k$. Saves us from numeric problems.

Programmed so far:

- constructing finite support prior
- splitting the space
- computing normalization factor

Prepared for:

- adaptivity – moving window – merging the space

Still needed:

- sampling from the distribution
- (maybe) computing moments

The model is much slower than the usual Gaussian type model. I've tested it for up to 10 parameters, where it gets quite slow. Moving window is needed. This might be ideal for financial

time series modelling, since if EMH holds, the model should be univariate and have only one scale and one location parameter. We can search the neighborhood for a better model.

Thank you for your attention.