High order low variance matrix-exponential distributions

Gábor Horváth[1] Illés Horváth[2] Miklós Telek[1]

[1]Budapest University of Technology and Economics, Hungary
[2]MTA-BME Information Systems Research Group, Hungary

MAM, Febr 13-15, 2019, Hobart, Australia

Outline

Long studied open problem:

Minimal CV of matrix exponential distributions
2006: conjectures and numerical values up to order 15
2016: minimal numerical values up to order 47
present: close to minimal up to order ~ 200
suboptimal values up to order ~ 1000

Matrix exponential (ME) distributions

Summary

- Distributions with density function

$$f(x) = \alpha e^{\boldsymbol{A}x}(-\boldsymbol{A})\mathbb{1},$$

  where α is a row vector, $\boldsymbol{A}$ is a square matrix and $\mathbb{1}$ is the column vector of ones
  no sign constraints, lack of stochastic interpretation.

- It is hard to check (in general) if $f(x)$ is non-negative.

Matrix exponential (ME) distributions

- The kth moment of ME with $f(x) = \alpha e^{\boldsymbol{A}x}(-\boldsymbol{A})\mathbb{1}$ is

$$\mu_k = \int_x x^k f(x)dx = k!\alpha(-\boldsymbol{A})^{-k}\mathbb{1},$$

- its squared coefficient of variation (SCV) is

$$SCV = \frac{\mu_0\mu_2}{\mu_1^2} - 1.$$

SCV is insensitive to multiplication and scaling, i.e. $SCV(f(x)) = SCV(c_1 f(c_2 x))$.

Phase type (PH) distributions

Summary

- Same as ME, but sign constraints apply for α and $\boldsymbol{A}$:
  - α is non-negative,
  - $\boldsymbol{A}$ has negative diagonal and non-negative off-diagonal elements.
- ME distributions of order N is a superset of phase type PH distributions of order N.

Properties

- Bounded SCV (Aldous-Shepp , O'Cinneide)

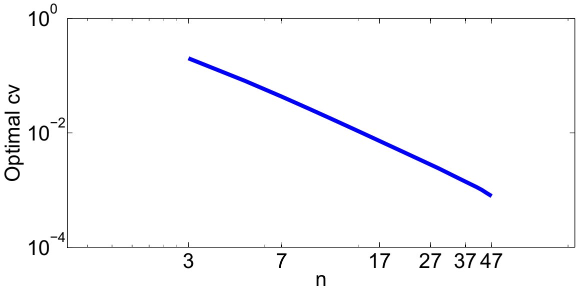
$$SCV \geq \frac{1}{N}.$$

Concentrated ME distributions

Only conjectures are available for minimal SCV of order N

N	SCV	$1/SCV$	date
3	0.20090	4.9776	2006
5	0.081264	12.306	
⋮	⋮	⋮	
15	0.0093128	107.38	
17	0.0072074	138.75	
19	0.0057368	174.31	2016
21	0.0046708	214.10	
⋮	⋮	⋮	
45	0.00088322	1132.2	
47	0.00078490	1274.0	

Concentrated ME distributions



- SCV decays with $2/N^2$ instead of $1/N$.

- This results are provided by a well defined ME "structure"
- using numerical optimization.

Exponential cosine-square functions

The set of *exponential cosine-square* functions of order n has the form

$$f^+(t) = ce^{-\eta t} \prod_{i=1}^{n} \cos^2\left(\frac{\omega t - \phi_i}{2}\right).$$

- $f^+(t)$ is non-negative by construction.
- Based on $SCV(f(x)) = SCV(c_1 f(c_2 x))$ we assume $c = \eta = 1$.
- $f^+(t)$ of order n is a subset of ME of order $N = 2n + 1$,
- with parameters ω and ϕ_i, $i = 1, \ldots, n$.

Persisting conjecture:
ME($2n+1$) with minimal SCV is $f^+(t)$ of order n.

Numerical optimization

Results for $N \leq 47$ are obtained by built in functions of general purpose (matlab, mathematica) packages.

For efficient numerical minimization of the SCV for $N > 47$ we need accurate and efficient computation methods with low computational complexity for

i) for computation of the SCV based on the parameters,

ii) for minimization of the SCV.

Hyper-trigonometric form of $f^+(t)$

Theorem

An exponential cosine-square function can be transformed to the following hyper-trigonometric form

$$f^+(t) = e^{-t} \prod_{i=1}^{n} \cos^2\left(\frac{\omega t - \phi_i}{2}\right)$$
$$= c^{(n)} \cdot e^{-t} + e^{-t} \sum_{k=1}^{n} a_k^{(n)} \cos(k\omega t)$$
$$+ e^{-t} \sum_{k=1}^{n} b_k^{(n)} \sin(k\omega t),$$

where $c^{(n)} = \frac{1}{2} a_0^{(n)}$ and the coefficients $a_k^{(n)}$, $b_k^{(n)}$ can be calculated recursively.

Hyper-trigonometric representation of $f^+(t)$

- for $n=1$: $a_0^{(1)}=1,\ b_0^{(1)}=0,\ a_1^{(1)}=\frac{1}{2}\cos\phi_1,\ b_1^{(1)}=\frac{1}{2}\sin\phi_1,$
- for $k>n, n\geq 1$: $a_k^{(n)}=b_k^{(n)}=0,$
- for $k=0, n\geq 1$:

$$a_0^{(n)}=\frac{1}{2}a_0^{(n-1)}+\frac{1}{2}a_1^{(n-1)}\cos\phi_n+\frac{1}{2}b_1^{(n-1)}\sin\phi_n,$$
$$b_0^{(n)}=0,$$

- for $1\leq k\leq n, n\geq 2$:

$$a_k^{(n)}=\frac{1}{2}a_k^{(n-1)}+\frac{1}{2}\frac{a_{k-1}^{(n-1)}+a_{k+1}^{(n-1)}}{2}\cos\phi_n+\frac{1}{2}\frac{b_{k+1}^{(n-1)}-b_{k-1}^{(n-1)}}{2}\sin\phi_n,$$
$$b_k^{(n)}=\frac{1}{2}b_k^{(n-1)}+\frac{1}{2}\frac{b_{k-1}^{(n-1)}+b_{k+1}^{(n-1)}}{2}\cos\phi_n+\frac{1}{2}\frac{a_{k-1}^{(n-1)}-a_{k+1}^{(n-1)}}{2}\sin\phi_n.$$

Hyper-trigonometric representation of $f^+(t)$

Based on the $a_k^{(n)}$, $b_k^{(n)}$ coefficients SCV is obtained efficiently.

Corollary

The $\mu_i, i = 0, 1, 2$ moments of the exponential cosine-square function are

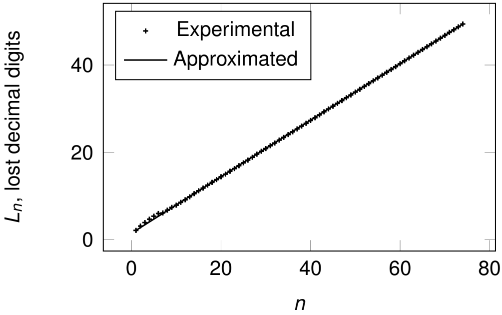
$$\mu_0 = c^{(n)} + \sum_{k=1}^{n} \frac{a_k^{(n)} + b_k^{(n)} k\omega}{1 + (k\omega)^2},$$

$$\mu_1 = c^{(n)} + \sum_{k=1}^{n} \frac{a_k^{(n)} + 2b_k^{(n)} k\omega - a_k^{(n)} (k\omega)^2}{(1 + (k\omega)^2)^2},$$

$$\mu_2 = 2c^{(n)} + \sum_{k=1}^{n} \frac{2a_k^{(n)} + 6b_k^{(n)} k\omega - 6a_k^{(n)} (k\omega)^2 - 2b_k^{(n)} (k\omega)^3}{(1 + (k\omega)^2)^3}.$$

Floting point computation of the coefficients

The computation of $a_k^{(n)}$, $b_k^{(n)}$ is prone to numerical errors



$$L_n \approx 1.487 + 0.647n.$$

The required floating point precision is $L_n + 16$ decimal digits to obtain an the coefficients in 16 decimal digits accuracy.

Floting point computation with the coefficients

Once the $a_k^{(n)}$, $b_k^{(n)}$ coefficients are computed with appropriate *high precision*,

standard precision is sufficient for all further computations.

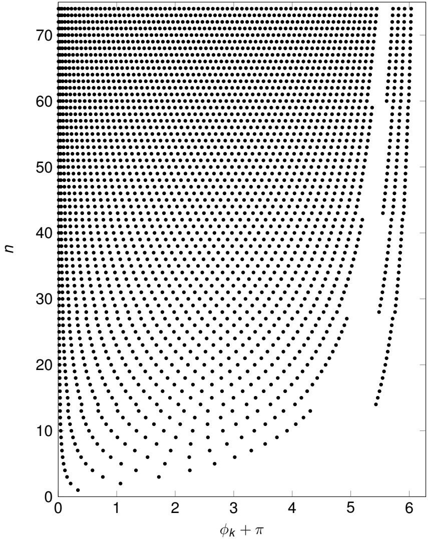
Optimization methods

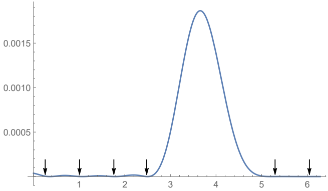
For finding the parameters (ω, ϕ_i) which provide the minimal SCV we got success with *evolution strategies*.

We compared three of them

$$T_{\text{CMA-ES}} < T_{\text{(1+1)-ES}} << T_{\text{BIPOP-CMA-ES}},$$
$$Q_{\text{CMA-ES}} \sim Q_{\text{(1+1)-ES}} < Q_{\text{BIPOP-CMA-ES}}.$$

Optimal parameters





Complexity as a function of n

The number of parameters to optimize is $n+1$,

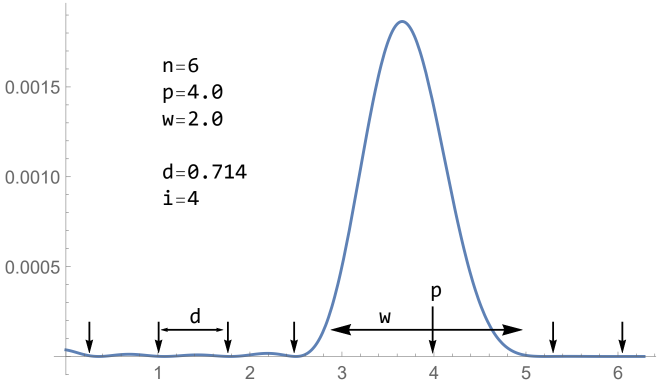
$\rightarrow$ complexity of optimization increases super linearly with n.

$\rightarrow$ complexity gets prohibitive around $n = 180$.

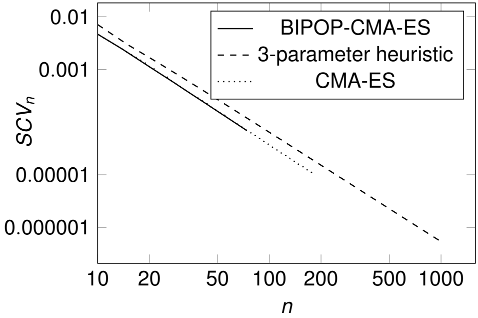
$\rightarrow$ low complexity suboptimal minimum is needed.

Heuristic approach with three parameters

- equidistant ϕ_i parameters, below and above the spike
- ϕ_i are defined by: place of spike (p), width of spike (w)



Heuristic approach



- The heuristic minimum is <2 times larger (for all n).
- ME distributions with SCV$<10^{-6}$ can be obtained.

Final remarks

Main messages:

- While the min SCV of *phase type* distribution decays *linearly*, **the min SCV of matrix exponential distribution decays quadratically** with the order.
- We can compute **concentrated matrix exponential distribution** up to order 10^3 **with SCV** $< \mathbf{10^{-6}}$.
- The obtained concentrated matrix exponential distribution can be efficiently used in stochastic models with deterministic delays with **fairly nice numerical stability** (next talk).