PIOTR GRACZYK
LAREMA, UNIVERSITÉ D'ANGERS, FRANCJA

Estymatory kowariancji dla modeli graficznych

współautorzy:
H. Ishi(Nagoya), B. Kołodziejek(Warszawa),
S. Mamane(Johannesburg), H. Ochiai(Fukuoka)

XLV Konferencja Zastosowań Matematyki,
Zakopane-Kościelisko, 9/9/2016

[GI] P. Graczyk, H. Ishi, *Riesz measures and Wishart laws associated to quadratic maps*, Journal of Jap.Math. Soc.66(2014), 317-348.

[GIM] P. Graczyk, H. Ishi, S. Mamane, *Wishart exponential families on cones related to A_n graphs*, submitted, 2016.

[GIK] P. Graczyk, H. Ishi, B. Kołodziejek, *Wishart exponential families and variance function on homogeneous cones*, submitted, 2016.

[GIMO] P. Graczyk, H. Ishi, S. Mamane, H. Ochiai, *On the Letac-Massam conjecture on cones Q_{A_n}*, submitted, 2016.

Objective:
Modern statistics and analysis of χ^2 laws on matrix cones

χ^2 laws on matrices are called Wishart laws

Let X be a Gaussian random vector $N(m, \Sigma)$ on $\mathbb{R}^r$, $m \in \mathbb{R}^r$, $\Sigma \in Sym_r^+$

$$f_X(x) = (\det 2\pi\Sigma)^{-1/2} \exp(-\frac{1}{2} {}^t(x-m)\Sigma^{-1}(x-m)).$$

X has unknown mean m and covariance Σ

We dispose of an observed sample $X^{(1)}, X^{(2)}, ..., X^{(s)}$

We want to estimate m and Σ, using the observed results $X^{(1)}, X^{(2)}, ..., X^{(s)}$

The sample mean

$$\bar{X}_s = \frac{X^{(1)} + X^{(2)} + ... + X^{(s)}}{s} = \widehat{m}$$

is an estimator of $m = \int X dP$

and the sample covariance

$$\widehat{\Sigma} = \frac{1}{s} \sum_{i=1}^{s} (X^{(i)} - \bar{X}_s)^t (X^{(i)} - \bar{X}_s)$$

is an estimator of $\Sigma = \int (X - m)^t (X - m) dP$.
$\widehat{\Sigma}$ is a quadratic map $\mathbb{R}^r \times \ldots \times \mathbb{R}^r \mapsto \overline{Sym^+(r, \mathbb{R})}$.

Proposition. The estimators $\widehat{m}$ and $\widehat{\Sigma}$ are **maximum likelihood estimators(MLE)**, i.e. they maximize the sample density treated as a function of parameters (*likelihood function*)

$$(m, \Sigma) \ \mapsto \ \prod_{i=1}^{s} f_X(x^{(i)}) = l(x^{(1)}, \ldots, x^{(s)}; m, \Sigma)$$

Proof. Maximizing of $\ln l(x^{(1)}, \ldots, x^{(s)}; m, \Sigma)$ in (m, Σ).

Modern research in this area? All seems known and done? NO!!!

Problems. Studying the MLE estimators for Σ when

1. Some entries of the vector X are known to be conditionally independent with respect to other ones.

It follows that Σ is submitted to some restrictive conditions; the range of Σ is no longer the whole cone $Sym^+(r, \mathbb{R})$

but a subcone $P \subset Sym^+(r, \mathbb{R})$ of matrices with obligatory zeros or its dual cone Q (explanation follows)

2. Some data is missing in observations $X^{(i)}$

A simplest example: We have s_1 "full" observations of X and s_2 observations of first k terms of X. How to estimate Σ from such a sample?

Conditional independence

Example: **Simpson paradox**

A university has 48 000 students
Half boys(24 000), half girls(24 000)

At the final exams: 10 000 boys and 14 000 girls fail
Feminist organizations threaten to close the university, girl students want to lynch the president!

However, the president of the university proves that the results R of the exams are independent of the sex S of a student, knowing the department D

3 departments:

A(sciences)

B(literature, history, languages)

C(law)

16 000 students each

A	Succ.	Fail	**B** Succ.	Fail	**C** Succ.	Fail
Girls	3	9	4	4	3	1
Boys	1	3	4	4	9	3

Actually $(R \perp S)|(D = d)$ for $d =$**A,B,C**

Conditional independence in a.c. case

$X = (X_1, X_2, X_3)$: Random vector

$f_{X_1,X_2,X_3}(x_1, x_2, x_3)$: density function

X_1 and X_3 are conditionally independent knowing X_2

$\Leftrightarrow f_{X_1,X_3|X_2=x_2} = f_{X_1|X_2=x_2} f_{X_3|X_2=x_2}$

$\Leftrightarrow f_{X_1,X_2,X_3}(x_1, x_2, x_3) = F(x_1, x_2) G(x_2, x_3)$

Example 1. $X \sim N(0, \Sigma), \quad \Sigma \in Sym_3^+$
$f_{X_1,X_2,X_3}(x_1, x_2, x_3) = (\det 2\pi\Sigma)^{-1/2} \exp(-{}^t x \Sigma^{-1} x/2)$

Put $\sigma := \Sigma^{-1}$. Mixing x_1 and x_3 can be avoided only when $\sigma_{13} = 0$:

$$f_{X_1,X_2,X_3}(x_1, x_2, x_3)$$
$$= (2\pi)^{-3/2}(\det \sigma)^{1/2} \exp\left(-(\sigma_{11}x_1^2 + 2\sigma_{12}x_1x_2 + \sigma_{22}x_2^2)/2\right)$$
$$\times \exp\left(-(2\sigma_{23}x_2x_3 + \sigma_{33}x_3^2)/2\right)$$

Therefore,
$(X_1 \perp X_3)|X_2 \;\Leftrightarrow\; \sigma_{13} = 0$
The matrix $\sigma = \Sigma^{-1}$ has obligatory zeros $\sigma_{13} = \sigma_{31} = 0$

The position of zeros in Σ^{-1} is encoded by a graph

$G = (V, E)$: undirected graph
$V = \{1, \ldots, r\}$: the set of vertices
$E \subset V \times V$: the set of edges
$i \sim j \Leftrightarrow (i, j) \in E$

$Z_G := \left\{ x \in \text{Sym}(r, \mathbb{R}) \,|\, x_{ij} = 0 \text{ if } i \neq j \text{ and } i \not\sim j \right\}$
$P_G := Z_G \cap Sym_r^+$ a sub-cone of Sym_r^+

$X \sim N(0, \Sigma)$, $\quad \Sigma^{-1} \in P_G$
$\Leftrightarrow$ X_i and X_j are conditionally independent knowing all other components if $i \neq j$ and $i \not\sim j$

Example 1 $(X_1 \perp X_3)|X_2$ corresponds to G: 1–2–3

Example 1. Graph $G = A_3$: 1–2–3

$$Z_G := \left\{ \begin{pmatrix} x_{11} & x_{12} & 0 \\ x_{12} & x_{22} & x_{23} \\ 0 & x_{23} & x_{33} \end{pmatrix} \mid x_{ij} \in \mathbb{R} \right\}$$

$P_G := Z_G \cap Sym_3^+$

This cone is homogeneous

($GL(P_G)$ acts transitively on P_G)

$$Z_G^* := \left\{ \begin{pmatrix} \xi_{11} & \xi_{12} & * \\ \xi_{12} & \xi_{22} & \xi_{23} \\ * & \xi_{23} & \xi_{33} \end{pmatrix} \,\middle|\, x_{ij} \in \mathbb{R} \right\}$$

$$\begin{aligned} P_G^* = Q_G &:= \left\{ \xi \in Z_G^* \,\middle|\, \mathrm{tr}\, x\xi > 0 \text{ for all } x \in \overline{\Omega_1} \setminus \{0\} \right\} \\ &= \left\{ \xi \in Z_G^* \,\middle|\, \begin{vmatrix} \xi_{11} & \xi_{12} \\ \xi_{12} & \xi_{22} \end{vmatrix} > 0,\ \begin{vmatrix} \xi_{22} & \xi_{23} \\ \xi_{23} & \xi_{33} \end{vmatrix} > 0,\ \xi_{33} > 0 \right\} \end{aligned}$$

Example 2. Graph $G = A_4$: 1–2–3–4

$$Z_G := \left\{ \begin{pmatrix} x_{11} & x_{21} & 0 & 0 \\ x_{21} & x_{22} & x_{32} & 0 \\ 0 & x_{32} & x_{33} & x_{43} \\ 0 & 0 & x_{43} & x_{44} \end{pmatrix} \,\middle|\, x_{11}, \ldots, x_{44} \in \mathbb{R} \right\}$$

$P_G := Z_G \cap S_4^+$

This cone is non-homogeneous

$$P_G^* = Q_G := \left\{ \xi \in Z_G^* \,|\, \mathrm{tr}\, x\xi > 0 \text{ for all } x \in \overline{\Omega_1} \setminus \{0\} \right\}$$

$$= \left\{ \xi \in Z_G^* \,\middle|\, \begin{vmatrix} \xi_{11} & \xi_{12} \\ \xi_{12} & \xi_{22} \end{vmatrix} > 0,\ \begin{vmatrix} \xi_{22} & \xi_{23} \\ \xi_{23} & \xi_{33} \end{vmatrix} > 0,\ \begin{vmatrix} \xi_{33} & \xi_{34} \\ \xi_{34} & \xi_{44} \end{vmatrix} > 0,\ \xi_{44} > 0 \right\}$$

Theory of graphical models
started in 1976 by Lauritzen and Speed,
is for decomposable graphs G

G is decomposable
$\Leftrightarrow G$ has no cycle of length ≥ 4 as an induced subgraph
Example: A_4 =1–2–3–4 from Example 2

$\Omega_G \subset Z_G$ is homogeneous if and only if
G is decomposable and A_4-free (Letac-Massam, Ishi)

Back to the estimation of the covariance matrix of a normal vector
Conditional independence case

Simplification:
$X = {}^{\mathrm{t}}(X_1, \ldots, X_n)$: random vector obeying $N(0, \Sigma)$
known mean 0, **unknown covariance matrix** $\Sigma \in S_n^+$
with $\Sigma^{-1} \in P_G = Z_G \cap S_n^+$

Sample $X^{(1)}, \ldots, X^{(s)} \in \mathbb{R}^n$

The MLE of Σ on the **whole** cone S_n^+ is (when $s \geq n$)

$$\tilde{\Sigma} = \frac{1}{s}\{X^{(1)}\,{}^{\mathrm{t}}X^{(1)} + X^{(2)}\,{}^{\mathrm{t}}X^{(2)} + \cdots + X^{(s)}\,{}^{\mathrm{t}}X^{(s)}\} \in \mathrm{Sym}(n, \mathbb{R})$$

Let us look for an estimator when $\Sigma^{-1} \in P_G$.

Example 1: G=1-2-3, $\sigma_{13} = 0$,
we shall estimate $\sigma = \Sigma^{-1} \in P_G$ and $\Sigma = \sigma^{-1}$

The density function of the sample $X^{(1)}, \ldots, X^{(s)} \in \mathbb{R}^3$
$f(x^{(1)}, \ldots, x^{(s)}; \sigma) =$
$= \prod_{k=1}^{s} \{(2\pi)^{-3/2} (\det \sigma)^{1/2} \exp(-{}^{\mathrm{t}}x^{(k)} \sigma x^{(k)}/2)\}$
$= (2\pi)^{-3s/2} (\det \sigma)^{s/2} \exp(-\sum_{k=1}^{s} {}^{\mathrm{t}}x^{(k)} \sigma x^{(k)}/2)$
Note that

$$\sum_{k=1}^{s} {}^{\mathrm{t}}x^{(k)} \sigma x^{(k)} = \mathrm{tr}\,(\sum_{k=1}^{s} x^{(k)}\,{}^{\mathrm{t}}x^{(k)})\sigma = \left\langle s\tilde{\Sigma}, \sigma \right\rangle = \left\langle \pi(s\tilde{\Sigma}), \sigma \right\rangle$$

where π = projection on Z_G^* (since $\sigma_{13} = 0$)

$$f(x^{(1)}, \dots, x^{(s)}; \sigma) = (2\pi)^{-3s/2} (\det \sigma)^{s/2} \exp(-\tfrac{1}{2} \left\langle \pi(s\tilde{\Sigma}), \sigma \right\rangle)$$

Which $\sigma \in P_G$ is **most likely?**

Maximum Likelihood Estimation $\Rightarrow$

it's $\sigma = \hat{\sigma}$ for which $f(x^{(1)}, \dots, x^{(s)}; \hat{\sigma})$ is maximum

We study $\log f(x^{(1)}, \dots, x^{(s)}; \sigma)$ as a function of $\sigma \in Z_G$

grad_σ does not contain $\frac{\partial}{\partial \sigma_{13}}$; $\quad \mathrm{grad} \log \det x = x^{-1}$

$$\mathrm{grad} \log f(x^{(1)}, \dots, x^{(s)}; \sigma) = \frac{s}{2}(\pi(\sigma^{-1}) - \pi(\tilde{\Sigma}))$$

If $s \geq 3$, then $\pi(\tilde{\Sigma})$ belongs to Q_G.

We must find $\widehat{\sigma} \in P_G$ such that

$$\pi(\widehat{\sigma}^{-1}) = \pi(\tilde{\Sigma}) \in Q_G$$

The inverse map Ψ to $x \in P_G \mapsto \pi(x^{-1}) \in Q_G$ is needed.
Only if $\pi = Id$ this is trivial ($\Psi(y) = y^{-1}$)

Then
$\widehat{\sigma} = \Psi(\pi(\tilde{\Sigma})) \in P_G$ is the MLE of σ
and, consequently
$\widehat{\Sigma} = \widehat{\sigma}^{-1} = (\Psi(\pi(\tilde{\Sigma}))^{-1}$ is the MLE of Σ.
$\pi(\tilde{\Sigma})$ is the MLE of $\pi(\Sigma)$, which determines Σ

The MLE argument is valid for any graphical cone.

CRUCIAL POINTS:

1. **law of** $\pi(\tilde{\Sigma})$ **on** Q_G

2. **knowledge of the inverse map** Ψ

law of $\pi(\tilde{\Sigma})$ on Q_G

$\pi(\tilde{\Sigma})$ is a quadratic map of the normal sample with values in Q_G:

$$\hat{\Sigma} = \pi(\tilde{\Sigma}) = \frac{1}{s}\pi(X^{(1)}\,{}^{\mathrm{t}}X^{(1)} + X^{(2)}\,{}^{\mathrm{t}}X^{(2)} + \cdots + X^{(s)}\,{}^{\mathrm{t}}X^{(s)})$$

$\pi(\tilde{\Sigma}) \in Q_G$ obeys a Wishart law on Q_G

Missing data

The simplest example:
$X =^t (X_1, X_2) \in \mathbb{R}^2$, $X \hookrightarrow N(\mu, \Sigma)$

We dispose of
s_1 "full" observations $X^{(1)}, \ldots, X^{(s_1)}$
s_2 incomplete observations $X'^{(1)}, \ldots, X'^{(s_2)}$
$X' =^t (X_1, *)$: the second entry X_2 is not observed

It is natural to set
$X' \hookrightarrow$ conditional law of $X_1|(X_2 = \mu_2) \quad = N(\mu_1, \frac{1}{\sigma_{11}})$,
$\sigma = \Sigma^{-1}$

$$f(x^{(1)},\dots,x^{(s_1)},x'^{(1)},\dots,x'^{(s_2)};\sigma) =$$
$$(2\pi)^{-(s_1+s_2)/2}\,(\det\sigma)^{s_1/2}\sigma_{11}^{s_2/2}\times$$
$$\exp\{-\tfrac{1}{2}[\textstyle\sum_i {}^t\bar{x}^{(i)}\sigma\bar{x}^{(i)} + \sum_j \sigma_{11}(\bar{x}'^{(j)})^2]\},\quad \bar{x} = x-\mu$$

The power function appears

$$\Delta_{(\frac{s_1+s_2}{2},\frac{s_1}{2})}(\sigma) = \sigma_{11}^{s_2/2}(\det\sigma)^{s_1/2}$$

The MLE equation $\mathrm{grad}_\sigma \log f = 0$ is

$$\mathrm{grad}\log\Delta_{(\frac{s_1+s_2}{2},\frac{s_1}{2})}(\sigma) = \pi(\tilde{q}(\bar{x}^{(1)},\dots,\bar{x}^{(s_1)},\bar{x}'^{(1)},\dots,\bar{x}'^{(s_2)}))$$

with a quadratic form $\tilde{q}(x)$ equal on the sample x to

$$\frac{1}{2}\sum_{i=1}^{s_1}\bar{x}^{(i)}\,{}^t\bar{x}^{(i)} + \begin{pmatrix}\frac{1}{2}\sum_{j=1}^{s_2}(\bar{x}'^{(j)})^2 & 0\\ 0 & 0\end{pmatrix} = q_{1,2}^{\otimes s_1}(x)\otimes q_1^{\otimes s_2}(x')$$

The inverse map $\Psi_{\underline{s}}$ to

$$x \in P_G \mapsto \text{grad} \log \Delta_{(\frac{s_1+s_2}{2}, \frac{s_1}{2})}(\sigma) \in Q_G$$

is needed.
We show its existence and write it explicitely ([GIM]).

Thus $\widehat{\sigma} = \Psi_{\underline{s}}(\pi(\tilde{q}(X)) \in P_G$ is the MLE of σ
and, consequently
$\pi(\tilde{q}(X))$ is the MLE of $\pi(\Sigma)$, which determines Σ

The equation

$$\pi(y^{-1}) = \xi \in Q_G, \quad y \in P_G$$

was solved by Lauritzen(1991) for decomposable graphs G, in terms of graph theory (cliques, separators...); the inverse map is called the **Lauritzen map**

We give such maps in monotonous missing data set-up:

$$\Psi_{\underline{s}} = (\pi(\mathrm{grad}\log\Delta_{(s_1,\ldots,s_n)})^{-1} = -\,\mathrm{grad}\log\delta_{-(s_1,\ldots,s_n)}$$

where $\delta_{-(s_1,\ldots,s_n)}$ is another power function (definitions later)

From now on,

$$G = A_n = 1 - 2 - \ldots - n$$

Q_{A_n} and P_{A_n} are important non-homogeneous($n \geq 4$) cones appearing in the statistical theory of graphical models

They correspond to the practical model of nearest neighbour interactions:

in the Gaussian character $(X_1, X_2, \ldots, X_n)$, non-neighbours X_i, X_j, $|i - j| > 1$ are conditionally independent with respect to other variables.

The Wishart laws on Q_G and P_G were first studied by Letac, Massam

G. Letac and H. Massam,
Wishart distributions for decomposable graphs,
The Annals of Statistics, **35** (2007), 1278–1323.

Motivations:

- **MLE theory**
Lauritzen, S.L. (1996) Graphical models. Oxford University Press

- **Bayesian statistics: searching conjugate prior distributions for $\pi(\Sigma) \in Q_G$ and $\sigma \in P_G$**
Diaconis P. and Ylvisaker D. (1979) Conjugate priors for exponential families. The Annals of Statistics 7(2):269-281.

Letac-Massam, following the Lauritzen graph theory methods, define power functions H on Q_G and P_G. For $G = A_n = 1 - 2 - \ldots - n$ they are defined:
on Q_{A_n} by:

$$H(\alpha, \beta, \eta) = \frac{\prod_{i=1}^{n-1} |\eta_{\{i:i+1\}}|^{\alpha_i}}{\prod_{i=2}^{n-1} \eta_{ii}^{\beta_i}}$$

and on P_{A_n} by

$$H(\alpha, \beta, \pi(y^{-1}))$$

Our results:

1. we introduce natural power functions

$$\delta_{\underline{s}}^{(M)}(\eta) \text{ on } Q_{A_n}, \quad \Delta_{\underline{s}}^{(M)}(y) \text{ on } P_{A_n}, \quad M = 1, \ldots, n$$

which contain (strictly) the Letac-Massam functions H. These functions are densities of **Riesz measures** on the cones Q_{A_n} and P_{A_n} respectively

2. We construct all Wishart families on Q_{A_n} and P_{A_n} generated by $\delta_{\underline{s}}^{(M)}(\eta)$ and $\Delta_{\underline{s}}^{(M)}(y)$. They contain (strictly) the Letac-Massam Wishart families $\gamma^{Q}_{(\alpha,\beta,y)}$ and $\gamma^{P}_{(\alpha,\beta,\eta)}$

We give their properties (density, moments)

3. we give an essential extension and simplification of Letac-Massam theory

4. we prove the Letac-Massam Conjecture on Q_{A_n}, on Laplace transform property of $H(\alpha, \beta, \eta)$

5. we find MLE of σ and $\pi(\Sigma)$ in the monotonous missing data case:

we construct an infinity of Lauritzen-type inverse maps $\Psi_{\underline{s}}: \ Q_{A_n} \mapsto P_{A_n}$

6. We determine the variance function $V(m)$ for the Wishart families on Q_{A_n}, where the mean $m = m_{\underline{s}}(y)$ is the expectation of $\gamma_{\underline{s},y}$.

In [GIK] $V(m)$ is determined for homogeneous cones

This is done thanks to an explicit form of the inverse mean map ψ_s.

$\Psi_{\underline{s}}$ and $\psi_{\underline{s}}$ are closely related: $\Psi_{\underline{s}} = -\psi_{-\underline{s}}$

Density of $\tilde{q}(X)$, X= a normal sample

In order to find MLEs, one needs quadratic forms $\tilde{q}(X^{(1)},\dots,X^{(s)}) \hookrightarrow$ Wishart law $\gamma_{\tilde{q},\sigma}$

the simplest case: $q(x) = x\,{}^t x$
$\mu_q = q(\text{Leb}_{\mathbb{R}^r})$ a Riesz measure

$\mathcal{L}_{\mu_q}(\eta) = \int_{\mathbb{R}^r} e^{-\langle \eta, x\,{}^t x\rangle} dx = \int_{\mathbb{R}^r} e^{-{}^t x \eta x} dx = \pi^{r/2}(\det \eta)^{-1/2}$
For an s-sample $\mathcal{L}_{\mu_q^{\oplus s}}(\eta) = \pi^{rs/2}(\det \eta)^{-s/2}$
For more general $\tilde{q}$: $\mathcal{L}_{\mu_{\tilde{q}}}(\eta) = c.\Delta_{-\underline{s}}(\eta)$

$$\Delta_{\underline{s}}(x) := \prod_{k=1}^{r} (\det x_{\{1,\dots,k\}})^{s_k - s_{k+1}}$$
$$= x_{11}^{s_1} \left(\frac{\det x_{\{1,2\}}}{x_{11}}\right)^{s_2} \cdots \left(\frac{\det x}{\det x_{\{1,\dots,r-1\}}}\right)^{s_r}$$

If μ is a measure on a cone $\Omega \subset V = \mathbb{R}^n$, then the family of probability measures

$$\gamma_y(dx) = \frac{e^{-(x,y)}}{\mathcal{L}(\mu)(y)}\mu(dx)$$

is called exponential family generated by μ.

The density of $\gamma_{\underline{s},\sigma}$ is $\frac{e^{-(y,\sigma)}}{\mathcal{L}_{\mu_{\underline{s}}}(\sigma)}\mu_{\underline{s}}(dy)$.
Thus, the density of the Riesz measure $\mu_{\underline{s}}$ is crucial.

When $\Omega = Sym_n^+ = S_n^+$, the density of the Riesz measure is given by the multiparameter Siegel integral:

For $\underline{s} \in \mathbb{C}^n$ with $\Re s_k > \frac{k-1}{2} \quad (k = 1, \ldots, n)$,

$$\int_{S_n^+} e^{-\mathrm{tr}\, x\xi} \delta_{\underline{s}}(x) (\det x)^{-\frac{n+1}{2}}\, dx = \Gamma_{S_n^+}(\underline{s}) \Delta_{-\underline{s}}(\xi) \quad (\xi \in S_n^+),$$

where

$$\delta_{\underline{s}}(x) = \left(\frac{\det x}{\det x_{\{2,\ldots,n\}}}\right)^{s_1} \ldots \left(\frac{\det x_{\{n-1,n\}}}{x_{nn}}\right)^{s_{n-1}} x_{nn}^{s_n},$$

$$\Gamma_{S_n^+}(\underline{s}) := \pi^{\frac{n-1}{2}} \prod_{k=1}^{n} \Gamma(s_k - \tfrac{k}{2})$$

POWER FUNCTIONS AND THEIR LAPLACE TRANSFORMS for $G = A_n$

Eliminating orders of vertices

There are many (but not all) orders of vertices $1, 2, \ldots, n$ that we should consider in order to have a harmonious theory of Riesz and Wishart distributions on the cones related to A_n graphs.

These orders are called *eliminating orders of vertices*.

Let v^+ be the set of future(w.r. to the order) neigbours (w.r. to the graph) of v.

An eliminating order of the vertices of G is a permutation $\{v_1, \ldots, v_n\}$ of V such that for all v, the set v^+ is a complete graph

Example. For the graph $A_3 : 1 - 2 - 3$:
the orders $1 \prec 2 \prec 3$, $1 \prec 3 \prec 2$, $3 \prec 2 \prec 1$ and $3 \prec 1 \prec 2$ are eliminating orders
$2 \prec 1 \prec 3$ and $2 \prec 3 \prec 1$ are not eliminating.

Proposition. All eliminating orders on A_n are obtained by an **intertwining of two sequences**
$1 \prec 2 \prec 3 < \ldots \prec M - 1 \prec M$
$n \prec n - 1 \prec \ldots \prec M + 2 \prec M + 1 \prec M$
for an $M \in V$.

Definition 1. *For* $\underline{s} \in \mathbb{C}^n$*, we define functions* $\Delta_{\underline{s}}^{\prec} : P_G \to \mathbb{C}^n$ *and* $\delta_{\underline{s}}^{\prec} : Q_G \to \mathbb{C}^n$ *by*

$$\Delta_{\underline{s}}^{\prec}(y) := \prod_{v \in V} \left(\frac{\det y_{\{v\} \cup v^-}}{\det y_{v^-}} \right)^{s_v} \qquad (y \in P_G),$$

$$\delta_{\underline{s}}^{\prec}(\eta) := \prod_{v \in V} \left(\frac{\det \eta_{\{v\} \cup v^+}}{\det \eta_{v^+}} \right)^{s_v} \qquad (\eta \in Q_G)$$

where $\det y_{\emptyset} = 1 = \det \eta_{\emptyset}$, $V = \{1, \ldots, n\}$

Property. On the cones Q_{A_n} and P_{A_n}, the power functions $\Delta_{\underline{s}}^{\prec}(y)$ and $\delta_{\underline{s}}^{\prec}(\eta)$ for eliminating orders $\prec$, depend only on the maximal element M of the order.
We write $\Delta_{\underline{s}}^{\prec}(y) = \Delta_{\underline{s}}^{(M)}(y)$, $\delta_{\underline{s}}^{\prec}(\eta) = \delta_{\underline{s}}^{(M)}(\eta)$

Proposition 2. *Let* $M = 1, \ldots, n$. *The "*M**-power functions"** $\Delta_{\underline{s}}^{(M)}(y)$ *on* P_G *and* $\delta_{\underline{s}}^{(M)}(x)$ *on* Q_G *are given by:*

$$\begin{aligned}\Delta_{\underline{s}}^{(M)}(y) = \; & y_{11}^{s_1 - s_2} \ldots |y_{\{1:M-1\}}|^{s_{M-1}-s_M} |y|^{s_M} \\ & \times |y_{\{M+1:n\}}|^{s_{M+1}-s_M} \ldots y_{nn}^{s_n - s_{n-1}}.\end{aligned}$$

$$\delta_{\underline{s}}^{(M)}(\eta) = \frac{\prod_{i=1}^{M-1} |\eta_{\{i:i+1\}}|^{s_i} \prod_{i=M+1}^{n} |\eta_{\{i-1:i\}}|^{s_i}}{\prod_{i=2}^{M-1} \eta_{ii}^{s_{i-1}} \cdot \eta_{MM}^{s_{M-1}-s_M+s_{M+1}} \cdot \prod_{i=M+1}^{n-1} \eta_{ii}^{s_{i+1}}}.$$

Theorem 3. *For all* $y \in P_{A_n}$,

$$\delta_{\underline{s}}^{(M)}(\pi(y^{-1})) = \Delta_{-\underline{s}}^{(M)}(y).$$

Characteristic function of a cone

$$\varphi_\Omega(x) = \int_{\Omega^*} e^{-(x,y)} dy = \mathcal{L}_{(\Omega^*, Leb)}(Leb_{\Omega^*})(x)$$

$\varphi_\Omega(x)dx$ is the invariant measure of the cone Ω:

$$\int_\Omega f(gx)\varphi_\Omega(x)dx = \int_\Omega f(x)\varphi_\Omega(x)dx.$$

For $n \geq 2$ define $\varphi_n : Q_{A_n} \to \mathbb{R}_+$ by

$$\varphi_n(\eta) = \prod_{i=1}^{n-1} |\eta_{\{i,i+1\}}|^{-3/2} \prod_{i \neq 1,n} \eta_{ii}$$

We will see that φ_n is the characteristic function of the cone Q_{A_n}.

Theorem 4. *For all* $n \geq 1$, *for all* $1 \leq M \leq n$ *and for all* $y \in P_{A_n}$, *the integral* $\int_{Q_{A_n}} e^{-\operatorname{tr}(y\eta)} \delta_{\underline{s}}^{(M)}(\eta) \varphi_{A_n}(\eta) d\eta$ *converges if and only if* $s_i > \frac{1}{2}$, *for all* $i \neq M$ *and* $s_M > 0$. *In this case, we have*

$$\begin{aligned} \int_{Q_{A_n}} & e^{-\operatorname{tr}(y\eta)} \delta_{\underline{s}}^{(M)}(\eta) \varphi_{A_n}(\eta) d\eta \\ = \quad & \pi^{(n-1)/2} \Big\{ \prod_{i \neq M} \Gamma(s_i - \frac{1}{2}) \Big\} \Gamma(s_M) \Delta_{-\underline{s}}^{(M)}(y). \end{aligned}$$

Theorem 5. *For all* $n \geq 1$, *for all* $1 \leq M \leq n$ *and for all* $\eta \in Q_{A_n}$, *the integral* $\int_{P_{A_n}} e^{-\mathrm{tr}(y\eta)} \Delta_{\underline{s}}^{(M)}(y)dy$ *converges if and only if* $s_i > -\frac{3}{2}$, *for all* $i \neq M$ *and* $s_M > -1$. *And in this case, we have*

$$\begin{aligned} &\int_{P_{A_n}} e^{-\mathrm{tr}(y\eta)} \Delta_{\underline{s}}^{(M)}(y)dy \\ = \;& \pi^{(n-1)/2} \Big\{ \prod_{i \neq M} \Gamma(s_i + \frac{3}{2}) \Big\} \Gamma(s_M + 1) \delta_{-\underline{s}}^{(M)}(\eta) \varphi_{A_n}(\eta). \end{aligned}$$

LETAC-MASSAM LAPLACE INTEGRALS

Recall Letac-Massam power functions on Q_{A_n}

$$H(\alpha,\beta,\eta)=\frac{\prod_{i=1}^{n-1}|\eta_{\{i:i+1\}}|^{\alpha_i}}{\prod_{i=2}^{n-1}\eta_{ii}^{\beta_i}}$$

The Laplace transform formula $\forall y\in P_{A_n}$

$$\int_{Q_{A_n}} e^{-\mathrm{tr}(y\eta)}H(\alpha,\beta,\eta)\varphi_{Q_{A_n}}(\eta)d\eta = C_{\alpha,\beta}H(\alpha,\beta,\pi^{-1}(y)),$$

will be referred to as the Letac-Massam formula on Q_{A_n}

Define $r_i = \alpha_i - \beta_{i+1}$, for all $1 \le i \le n-3$ and $p_i = \alpha_i - \beta_i$, for all $3 \le i \le n-1$. We have

$$H(\alpha, \beta, \eta) = \delta_{\underline{s}}^{(M)}(\eta) \prod_{i=2}^{M-1} \eta_{ii}^{r_{i-1}} \prod_{i=M+1}^{n-1} \eta_{ii}^{p_i},$$

where $s_i = \alpha_i$, for all $1 \le i \le M-1$; $s_i = \alpha_{i-1}$, for all $M+1 \le i \le n$ and $\beta_M = s_{M-1} - s_M + s_{M+1}$.

We have proved

Theorem. The Letac-Massam formula on Q_{A_n} holds if and only if

$$H(\alpha, \beta, \eta) = \delta_{\underline{s}}^{(M)}(\eta)$$

for some $M = 2, \ldots, n-1$.

(a new formulation of "Letac-Massam conjecture")

Methods. Change of variables

Let $\Phi_n : \mathbb{R}^+ \times \mathbb{R} \times P_{A_{n-1}} \longrightarrow P_{A_n}$, $(a, b, z) \longmapsto y$ with

$$y = \begin{pmatrix} 1 & & & & \\ b & & \ddots & & \\ 0 & & & & \\ \vdots & & & \ddots & \\ 0 & \dots\dots\dots & & 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & z & \\ 0 & & & \end{pmatrix} \begin{pmatrix} 1 & & & & \\ b & & \ddots & & \\ 0 & & & & \\ \vdots & & & \ddots & \\ 0 & \dots\dots\dots & & 0 & 1 \end{pmatrix}^T$$

Let $\Psi_n : \mathbb{R}^+ \times \mathbb{R} \times Q_{A_{n-1}} \longrightarrow Q_{A_n}$, $(\alpha, \beta, x) \longmapsto \eta$ with

$$\eta = \pi \left(\begin{pmatrix} 1 & & & & \\ \beta & & \ddots & & \\ 0 & & & & \\ \vdots & & & \ddots & \\ 0 & \dots\dots\dots & & 0 & 1 \end{pmatrix}^T \begin{pmatrix} \alpha & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & x & \\ 0 & & & \end{pmatrix} \begin{pmatrix} 1 & & & & \\ \beta & & \ddots & & \\ 0 & & & & \\ \vdots & & & \ddots & \\ 0 & \dots\dots\dots & & 0 & 1 \end{pmatrix} \right)$$

The maps Φ_n and Ψ_n are bijections.

Back to MLE of σ and Σ

Theorem The map

$$-\text{grad}\,\log\Delta^{(M)}_{-\underline{s}} : P_{A_n} \to Q_{A_n}$$

has the inverse map

$$\text{grad}\,\log\delta^{(M)}_{\underline{s}} : Q_{A_n} \to P_{A_n}$$

proved on homogeneous cones by Kai-Nomura (2005)

$\Rightarrow$ solution of MLE equation

$$\text{grad}\,\log\Delta_{\underline{s}}(\widehat{\sigma}) = \pi(q_{1,2}^{\otimes s_1}(x) \otimes q_1^{\otimes s_2}(x'))$$