The Gaussian Correlation Conjecture: revision of a proof by T. Royen

Raúl Astudillo Marbán

Universidad de Guanajuato/CIMAT

12/21/2015

Introduction

Conjecture A (Gaussian Correlation Conjecture). Let $\mathbb{P}$ be a probability measure in $\mathbb{R}^n$ given by a Gaussian density:

$$f(x) = \frac{1}{(2\pi)^{\frac{n}{2}} \left|\Sigma\right|^{\frac{1}{2}}} e^{-\frac{1}{2}x^T\Sigma^{-1}x}, x \in \mathbb{R}^n$$

whit a non-singular covariance matrix Σ. If A and B are closed, convex and symmetric about the origin, then

$$\mathbb{P}(A \cap B) \geq \mathbb{P}(A)\mathbb{P}(B).$$

Preliminary results

Conjecture B. If $(X_1, ..., X_n)$ is a $N_n(0, \Sigma)$-Gaussian vector (Σ not necessarily non-singular), then for any $x_1, ..., x_n \geq 0$,

$$\mathbb{P}(\bigcap_{i=1}^{n} A_i) \geq \mathbb{P}(\bigcap_{i=1}^{k} A_i)\mathbb{P}(\bigcap_{i=k+1}^{n} A_i),$$

where $1 \leq k < n$, $A_i = \{|X_i| \leq x_i\}$, $i = 1, ..., n$.

Theorem. Conjectures A and B are equivalent.

Proposition 1. Let f be the $\Gamma_n(\alpha, R)$ probability density function with $2\alpha \in \mathbb{N}$. If R is represented by

$$R = \lambda I_n + AA^T,$$

where λ is the minimal eigenvalue of R and A is $n \times (n-1)$ matrix, then

$$f(x_1, ..., x_n; \alpha, R) = \mathbb{E}\left(\prod_{j=1}^{n} \lambda^{-1} g_\alpha(\lambda^{-1} x_j, \frac{1}{2} b_j S b_j^T)\right),$$

where $g_\alpha(x, y)$ is the non-central gamma probability density function, b_j, $j = 1, ..., n$, are the rows of $B = \lambda^{-\frac{1}{2}} A$ and S is a $W_{n-1}(2\alpha, I_{n-1})-$Wishart matrix.

This formula can be found in a more general form in

Royen, T. (2007). Integral representations and approximations for multivariate gamma distributions, *Ann. Inst. Statist. Math.* 59, 499–513.

Proposition 2. Let $R = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix}$ be a non-singular correlation matrix and for every $t \in [0, 1]$ define the matrix

$$R_t = \begin{pmatrix} R_{11} & tR_{12} \\ tR_{21} & R_{22} \end{pmatrix}$$

R_t is symmetric and positive definite. In particular, if λ_t is the minimal eigenvalue of R_t, then $\lambda_t > 0$.

Proof. Since $R_1 = R$, it is positive definite. Now let n_1 be the dimension of R_{11}. For $x = (x_1, ..., x_{n_1}, x_{n_1+1}, ..., x_n) \in \mathbb{R}^n \backslash \{0\}$ we have

$$\begin{aligned} xR_0x^T &= x \begin{pmatrix} R_{11} & 0 \\ 0 & R_{22} \end{pmatrix} x^T \\ &= x_1 R x_1^T + x_2 R x_2^T \\ &> 0 \end{aligned}$$

where $x_1 = (x_1, ..., x_{n_1}, 0, ..., 0)$, $x_2 = (0, ..., 0, x_{n_1+1}, ..., x_n)$.

Thus, R_0 is also positive definite.

For every $t \in [0,1]$ R_t can be represented as a convex combination of two positive definite matrices:

$$R_t = tR_1 + (1-t)R_0,$$

so all of them are positive definite.

Main theorem

Theorem. Let R and R_t be defined as in the proposition above. For any $x_1, ..., x_n \geq 0$ and $\alpha > 0$ such that $2\alpha \in \mathbb{N}$ the function $t \mapsto F(x_1, ..., x_n; \alpha, R_t)$, where $F(x_1, ..., x_n; \alpha, R_t)$ is the $\Gamma_n(\alpha, R_t)$-cumulative distribution function, is increasing in $[0, 1]$.

In particular,

$$F(x_1, ..., x_n; \alpha, R) \geq F(x_1, ..., x_{n_1}; \alpha, R_{11}) F(x_{n_1+1}, ..., x_n; \alpha, R_{22}).$$

Using this theorem we can easily prove the classical Gaussian Correlation Conjecture.

Theorem (Gaussian Correlation Conjecture). If $(X_1, ..., X_n)$ is a $N_n(0, \Sigma)$-Gaussian vector (Σ not necessarily non-singular), then for any $x_1, ..., x_n \geq 0$,

$$\mathbb{P}(\bigcap_{i=1}^{n} A_i) \geq \mathbb{P}(\bigcap_{i=1}^{k} A_i)\mathbb{P}(\bigcap_{i=k+1}^{n} A_i),$$

where $1 \leq k < n$, $A_i = \{|X_i| \leq x_i\}$, $i = 1, ..., n$.

Proof. Without loss of generality we can assume that $\alpha_i = Var(X_i) > 0$, $i = 1, ..., n$.

The vector $(\frac{X_1^2}{\alpha_1}, ..., \frac{X_n^2}{\alpha_n})$ is $\Gamma_n(\frac{1}{2}, R)$-distributed for some correlation matrix R.

As before we write R as a partitioned matrix:

$$R = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix},$$

where R_{11} is a $k \times k$ matrix.

We see that

$$A_i = \{|X_i| \leq x_i\} = \{\frac{X_i^2}{\alpha_i} \leq \frac{x_i^2}{\alpha_i}\}.$$

Hence,

$$\mathbb{P}(\bigcap_{i=1}^{n} A_i) = F(\frac{x_1^2}{\alpha_1}, ..., \frac{x_n^2}{\alpha_n}; \frac{1}{2}, R),$$

while

$$\mathbb{P}(\bigcap_{i=1}^{k} A_i) = F(\frac{x_1^2}{\alpha_1}, ..., \frac{x_k^2}{\alpha_k}; \frac{1}{2}, R_{11}),$$

$$\mathbb{P}(\bigcap_{i=k+1}^{n} A_i) = F(\frac{x_{k+1}^2}{\alpha_{k+1}}, ..., \frac{x_n^2}{\alpha_n}; \frac{1}{2}, R_{22}).$$

From the theorem above it follows that

$$\mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right) \geq \mathbb{P}\left(\bigcap_{i=1}^{k} A_i\right)\mathbb{P}\left(\bigcap_{i=k+1}^{n} A_i\right). \blacksquare$$

Proof of the main theorem

It would be enough to prove that

$$\frac{d}{dt}F(x_1,...,x_n;\alpha,R_t) > 0 \text{ for } t \in (0,1).$$

For a $n \times n$ matrix A let A_J be the submatrix with row and column indices $i \in J \subseteq \{1, ..., n\}$.

It is not hard to see that

$$|I_n + RT| = 1 + \sum_{J \subseteq \{1,...,n\}} |R_J| \, |T_J|$$

where $T = diag(t_1, ..., t_n)$.

For every $J \subseteq \{1, ..., n\}$ with $J_1 = J \cap \{1, ..., n_1\} \neq \emptyset$, $J_2 = J \cap \{n_1 + 1, ..., n\} \neq \emptyset$ and $t \in [0, 1]$ define the matrix

$$R_{t,J} = \begin{pmatrix} R_{J_1} & tR_{J_1,J_2} \\ tR_{J_2,J_1} & R_{J_2} \end{pmatrix}.$$

Since R is a correlation matrix, R_{J_1}, R_{J_2} and $R_{t,J}$, $t \in [0, 1]$ are all symmetric and positive definite.

Now let r_{J_1,J_2} =rank(R_{J_1,J_2}) and $\rho^2_{J_1,J_2,i}$, $i = 1, ..., r_{J_1,J_2}$ the canonical correlations, which are the positive eigenvalues of

$$R_{J_1}^{-\frac{1}{2}} R_{J_1,J_2} R_{J_2}^{-1} R_{J_2,J_1} R_{J_1}^{-\frac{1}{2}}.$$

Since $R_{1,J}$ is non-singular, $0 < \rho^2_{J_1,J_2,i} < 1$, $i = 1, ..., r_{J_1,J_2}$.

Using the formula for the determinant of a partitioned matrix

$$
\begin{aligned}
|R_{t,J}| &= |R_{J_1}|\,|R_{J_2}|\left|I_{|J_1|} - t^2 R_{J_1}^{\frac{1}{2}} R_{J_1,J_2} R_{J_2}^{-1} R_{J_2,J_1} R_{J_1}^{-\frac{1}{2}}\right| \\
&= |R_{J_1}|\,|R_{J_2}| \prod_{i=1}^{r_{J_1,J_2}} (1 - t^2 \rho_{J_1,J_2,i}^2).
\end{aligned}
$$

It follows that

$$\begin{aligned}
\frac{d}{dt}\left|R_{t,J}\right| &= \left|R_{J_1}\right|\left|R_{J_1}\right|\frac{d}{dt}\prod_{i=1}^{r_{J_1,J_2}}(1-t^2\rho^2_{J_1,J_2,i}) \\
&= -\left|R_{J_1}\right|\left|R_{J_1}\right|\sum_{i=1}^{r_{J_1,J_2}}\left(2t\rho^2_{J_1,J_2,i}\prod_{j\neq i}(1-t^2\rho^2_{J_1,J_2,i})\right) \\
&= -2t\left|R_{J_1}\right|\left|R_{J_1}\right|\prod_{i=1}^{r_{J_1,J_2}}(1-t^2\rho^2_{J_1,J_2,i})\sum_{i=1}^{r_{J_1,J_2}}\frac{\rho^2_{J_1,J_2,i}}{1-t^2\rho^2_{J_1,J_2,i}} \\
&= -2t\left|R_{t,J}\right|\sum_{i=1}^{r_{J_1,J_2}}\frac{\rho^2_{J_1,J_2,i}}{1-t^2\rho^2_{J_1,J_2,i}}.
\end{aligned}$$

Using the identity above we get

$$
\begin{aligned}
& \frac{d}{dt}\left|I_n + R_t T\right|^{-\alpha} \\
= & -\alpha \left|I_n + R_t T\right|^{-\alpha-1} \frac{d}{dt}\left|I_n + R_t T\right| \\
= & -\alpha \left|I_n + R_t\right|^{-\alpha-1} \frac{d}{dt}\left(1 + \sum_J \left|R_{t,J}\right| \left|T_J\right|\right) \\
= & -\alpha \left|I_n + R_t\right|^{-\alpha-1} \left(\sum_J -2t\left|R_{t,J}\right| \left(\sum_{i=1}^{r_{J_1,J_2}} \frac{\rho^2_{J_1,J_2,i}}{1 - t^2\rho^2_{J_1,J_2,i}}\right) \prod_{j\in J} t_j\right) \\
= & \; 2\alpha t \left|I_n + R_t\right|^{-\alpha-1} \sum_J \left|R_{t,J}\right| \left(\sum_{i=1}^{r_{J_1,J_2}} \frac{\rho^2_{J_1,J_2,i}}{1 - t^2\rho^2_{J_1,J_2,i}}\right) \prod_{j\in J} t_j \\
= & \left|I_n + R_t\right|^{-\alpha-1} \sum_J c_J(t) \prod_{j\in J} t_j,
\end{aligned}
$$

where

$$c_J(t) = 2\alpha t \, |R_{t,J}| \left(\sum_{i=1}^{r_{J_1,J_2}} \frac{\rho^2_{J_1,J_2,i}}{1 - t^2 \rho^2_{J_1,J_2,i}} \right).$$

Now define

$$h^*(t_1, ..., t_n; \alpha, R_t) = |I_n + R_t|^{-\alpha-1} \sum_J c_J(t) \prod_{j \in J} t_j,$$

$$h(x_1, ..., x_n; \alpha, R_t) = \sum_J c_J(t) \left(\prod_{j \in J} \frac{\partial}{\partial x_j} \right) f(x_1, ..., x_n; \alpha + 1, R_t).$$

It is not hard to see that h^* is the Laplace transform of h.

We now want to prove that

$$h(x_1, ..., x_n; \alpha, R_t) = \frac{d}{dt} f(x_1, ..., x_n; \alpha, R_t).$$

We see that

$$\begin{aligned}
&\int_{\mathbb{R}_+^n} \left(f(x_1, ..., x_n; \alpha, R_t) - f(x_1, ..., x_n; \alpha, R_0)\right) \prod_{j=1}^{n} e^{-t_j x_j} dx \\
= \;& |I_n + R_t T|^{-\alpha} - |I_n + R_0 T|^{-\alpha} \\
= \;& \int_0^t h^*(t_1, ..., t_n; \alpha, R_s) ds \\
= \;& \int_0^t \int_{\mathbb{R}_+^n} h(x_1, ..., x_n; \alpha, R_s) \prod_{j=1}^{n} e^{-t_j x_j} dx ds \\
= \;& \int_{\mathbb{R}_+^n} \left(\int_0^t h(x_1, ..., x_n; \alpha, R_s) ds \right) \prod_{j=1}^{n} e^{-t_j x_j} dx.
\end{aligned}$$

It would be enough to justify the change in the order of integration in the last equality.

$$
\begin{aligned}
& \left|h(x_1, \ldots, x_n; \alpha, R_s)\right| \\
\leq & \sum_J c_J(s) \left| \left(\prod_{j \in J} \frac{\partial}{\partial x_j} \right) f(x_1, \ldots, x_n; \alpha + 1, R_s) \right| \\
= & \sum_J c_J(s) \left| \left(\prod_{j \in J} \frac{\partial}{\partial x_j} \right) E \left[\prod_{j=1}^{n} \lambda_s^{-1} g_{\alpha+1}(\lambda_s^{-1} x_j, \frac{1}{2} b_j S b_j^T) \right] \right| \\
= & \sum_J c_J(s) \left| E \left[\left(\prod_{j \in J} \frac{\partial}{\partial x_j} \right) \prod_{j=1}^{n} \lambda_s^{-1} g_{\alpha+1}(\lambda_s^{-1} x_j, \frac{1}{2} b_j S b_j^T) \right] \right| .
\end{aligned}
$$

Applying the identity

$$\frac{\partial}{\partial x} g_{\alpha+1}(x,y) = g_\alpha(x,y) - g_{\alpha+1}(x,y)$$

we get

$$\begin{aligned}
& \sum_J c_J(s) \left| E\left[\left(\prod_{j\in J} \frac{\partial}{\partial x_j} \right) \prod_{j=1}^n \lambda_s^{-1} g_{\alpha+1}(\lambda_s^{-1} x_j, \frac{1}{2} b_j S b_j^T) \right] \right| \\
= \; & \sum_J c_J(s) \lambda_s^{-|J|} \left| E\left[\prod_{j=1}^n \lambda_s^{-1} \left(e_J(j) g_\alpha(\lambda_s^{-1} x_j, \frac{1}{2} b_j S b_j^T) - g_{\alpha+1}(\lambda_s^{-1} x_j, \frac{1}{2} b_j S b_j^T) \right) \right] \right| \\
\leq \; & \sum_J c_J(s) \lambda_s^{-|J|} E\left[\prod_{j=1}^n \lambda_s^{-1} \left(e_J(j) g_\alpha(\lambda_s^{-1} x_j, \frac{1}{2} b_j S b_j^T) + g_{\alpha+1}(\lambda_s^{-1} x_j, \frac{1}{2} b_j S b_j^T) \right) \right].
\end{aligned}$$

Therefore

$$\int\limits_{\mathbb{R}_+^n} |h(x_1, ..., x_n; \alpha, R_s)| \prod_{j=1}^{n} e^{-t_j x_j} dx$$

is bounded by a finite linear combination of integrals of the form

$$\begin{aligned} & \int\limits_{\mathbb{R}_+^n} E\left[\prod_{j=1}^{n} \lambda_s^{-1} g_{\alpha+e_K(j)}(\lambda_s^{-1} x_j, \frac{1}{2} b_j S b_j^T)\right] \prod_{j=1}^{n} e^{-t_j x_j} dx \\ = \quad & |I_n + R_s T|^{-\alpha-1} \prod_{i \notin K} (1 + \lambda_s t_i) \end{aligned}$$

where $K \subseteq \{1, ..., n\}$, $K \neq \emptyset$. The coefficients of this linear combination are functions of the form $c_J(s)\lambda_s^{-|J|}$, which are continuous and non-negative in $[0, 1]$.

This implies that

$$\int_0^t \int_{\mathbb{R}_+^n} |h(x_1, ..., x_n; \alpha, R_s)| \prod_{j=1}^n e^{-t_j x_j} dx < \infty,$$

and thus the change in the order of integration is justified by Fubini's theorem.

Hence we have proved that

$$\begin{aligned}\frac{d}{dt} f(x_1, ..., x_n; \alpha, R_t) &= h(x_1, ..., x_n; \alpha, R_t) \\ &= \sum_J c_J(t) \left(\prod_{j \in J} \frac{\partial}{\partial x_j} \right) f(x_1, ..., x_n; \alpha + 1, R_t).\end{aligned}$$

Integration over $x_1, ..., x_n$ leads to

$$\frac{d}{dt}F(x_1, ..., x_n; \alpha, R_t) = \sum_J c_J(t) \left(\prod_{j \in J} \frac{\partial}{\partial x_j} \right) F(x_1, ..., x_n; \alpha+1, R_t) \geq 0.$$

This finishes the proof.

Remark: by continuity the inequality also holds for R singular.

References

1. Royen, T. (2014). A simple proof of the Gaussian correlation conjecture extended to multivariate gamma distributions, *Far East Journal of Theoretical Statistics.* 48, 139-145.
2. Royen, T. (2007). Integral representations and approximations for multivariate gamma distributions, *Ann. Inst. Statist. Math.* 59, 499–513.