Leibniz FH

Research Paper Series
Nr. 7 | March 2020

A Note on the Geometry of Partial Correlation and the Grassmannian

Autor:
Roland Matthes

Leibniz-Fachhochschule

Autor:

Roland Matthes
Leibniz-Fachhochschule
matthes@leibniz-fh.de

ISSN 2511-7491

Redaktion: Robin Christmann

Leibniz FH
School of Business

Research Paper No. 7

A Note on the Geometry of Partial Correlation and the Grassmannian

ROLAND MATTHES

Zusammenfassung / Abstract:

In this paper partial correlation is explained geometrically as the angle between certain linear subspaces. They can be computed from an inner product between the corresponding pure vectors on the exterior algebra.

Fachrichtung / Field of Study: Mathematics

Klassifikation / Classification: 62H15, 6H20, 14M15, 15A75

(Mathematics Subject Classification 2010)

Schlagworte / Keywords: -

1 Introduction

Without doubt in all branches of mathematics geometric arguments, if available, provide a source for intuition and may help for a better understanding of underlying principles. A prominent exampe in statistics is given by the covariance and correlation of two random variables. Covariance can be understood as an inner product, so it inherits a geometric meaning and correlation thus can be interpreted as (the cosine of) an angle between these variables.

The empirical counterpart, the sample correlation coefficient, then corresponds to the angle between the standardized vectors of observations (sample vectors).

For the partial correlation coefficient $\rho_{12;3..m}$in (2), a geometric interpretation in terms of spherical geometry can be found in [10], [1] for $m = 3$.

In our present paper we shall extend this interpretation to arbitrary $m \geq 3$ and express partial correlation coefficient in terms of an inner product between flats that are spanned by the involved standardized sample vectors.

To this end we will consider k-dimensional subspaces (flats containing the origin) of $\mathbb{R}^n$ as elements of the Grassmannian $Gr(k, n)$ which can be employed with the structure of a Riemannian space, where the inner product on the tangent space is inherited from the Euclidean structure of the exterior product $\Lambda^k(\mathbb{R}^n)$. The latter is just the pullback of the canonical inner product on $\mathbb{R}^{\binom{n}{k}} \cong \Lambda^k(\mathbb{R}^n)$ under the Plücker embedding..

This inner product is known to induce the Fubini-Study metric on the Grassmannian and it can be expressed as the product of the cosines of the principal angles between the flats.

2 Partial correlation

If one is interested whether there is a (possibly linear) dependence between two random variables X^1, X^2 the sample correlation coefficients may serve as a reasonable statistical entity to measure such dependence. Of course statistical evidence does not imply a causal relationship. There may exist further external random variables $X^3, ..., X^m$ which exercise influence on both X^1 and X^2 and therefore a statistical correlation between X^1 and X^2 would be caused by external effects.

Partial correlation between X^1 and X^2 with respect to $X^3, ..., X^m$ is intended to measure the 'true' correlation, meaning to remove the effect of the external variables.

2.1 Linear regression approach

We always have to distingish between the theoretical and the empirical entities. The theoretical correlation coefficient ρ_{ij} between X^i and X^j is as usual given by the ratio $\frac{Cov(X^i,X^j)}{\sqrt{Var(X^i)Var(X^j)}}$. We now fix some basic notations for the definition of the sample correlation coefficient. For a (real valued) sample $x^i = (x^i_1, ..., x^i_n)^T$ of size n concerning the random variable X^i we define the usual sample mean

$$\bar{x^i} = \frac{1}{n}\sum_{k=1}^{n} x^i_k$$

the sample variance

$$s_{x^i}^2 = \frac{1}{n-1}\sum_{k=1}^{n}(x_k^i - \bar{x})^2$$

and the standardized sample variable

$$\xi_k^i = \frac{x_k^i - \bar{x}}{\sqrt{n-1}\, s_{x^i}}. \tag{1}$$

The sample correlation coefficient between samples x^i and x^j due to Pearson is then given by the inner product

$$r_{ij} = \sum_{k=1}^{n} \xi_k^i \xi_k^j = \langle \xi^i, \xi^j \rangle.$$

If $X^1 \sim a + \sum_{i=3} b_i X^i$ and $X^2 \sim c + \sum_{i=3} d_i X^i$ are best linear approximation models based on samples x^i of n observations for X^i, then theoretical partial correlation is defined to be the correlation coefficient between $X^1 - (a + \sum_{i=3}^{n} b_i X^i)$ and $X^2 - (c + \sum_{i=3}^{n} d_i X^i)$.

Explicitly it is given by

$$\rho_{12;3..m} = -\frac{P_{12}}{\sqrt{P_{11}P_{22}}} \tag{2}$$

where P_{ij} denotes the cofactor in the determinant of the matrix $P = (\rho_{ij})$ of the correlation coefficients between X^i and X^j. (So it is the determinant of P with the i-th row and j-th column removed and multiplied by $(-1)^{i+j}$.) See e.g. [9], ch. 28.

The empirical analogue is the sample partial correlation coefficient

$$r_{12;3..m} = -\frac{R_{12}}{\sqrt{R_{11}R_{22}}} \tag{3}$$

and R_{ij} denotes the cofactor in the determinant of the matrix $R = (r_{ij})$

2.2 Geometric approach

Another way to descibe the influence of the external variates is geometrical in nature. First let us consider the case of one additional external variable X^3 with sample vector ξ^3.

If $n = 3$ it is shown in [10], that the sample partial correlation coefficient corresponds to the cosine of the spherical angle $\theta_{spherical}$ between ξ^1 and ξ^2 when watched form ξ^3

$$r_{12;3} = cos(\theta_{sperical}). \tag{4}$$

see figure 2.2. In 3-dimensional spherical geometry the angle $\theta_{sperical}$ is defined as the angle between the planes that intersect the surface of the plane, so in this case it is the angle between $L_1 = L(\xi^1, \xi^3)$ and $L_2 = L(\xi^2, \xi^3)$.

If ξ^3 was perpendicular to both ξ^1 and ξ^2,i.e.$r_{13} = r_{23} = 0$, then it is obvious, that the angle between the vectors ξ^1 and ξ^2, is the same as the spherical angle.

If in the more or less opposite case ξ^3 is a linear combination of ξ^1 and ξ^2 then viewed from ξ^3 the other sample variables are in perfect correlation. This corresponds to the fact, that the angle between the $L1$ and $L2$ is zero.

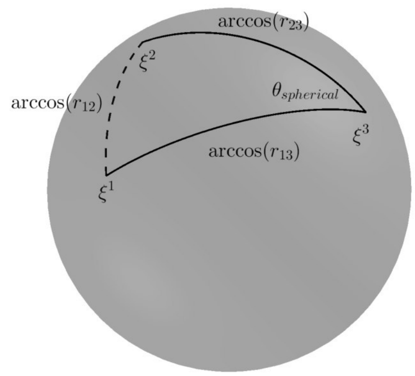

Figure 1: $r_{12;3} = \cos(\theta_{spherical})$

The above formula (4) generalizes for $m = 3$ to n dimensions, when projecting $\mathbb{R}^n$ to $L(\xi^1, \xi^2, \xi^3)$. So also in this case, the partial correlation coefficient is given by the cosine of the angle between the planes L_1 and L_2.

In the next section we show, how this can be generalized within a suitable geometric framework to the case of more than one external variable, i.e. $m > 3$.

3 Grassmannian

When we are concerned with several external variables, we have to consider angles between k-dimensional linear subspaces of $\mathbb{R}^n$. When working with such flats as objects it is convenient to regard these flats as points on the above mentionend manifold $Gr(k, n)$ the Grassmannian, which as a set can be identified with a subset of the projective exterior product $\mathbb{P}(\Lambda^k(\mathbb{R}^n))$.

We review the most basic facts about the exterior product.

3.1 Exterior product

The k-th exterior power $\Lambda^k(\mathbb{R}^n)$ consisting of the so-called k-vectors is an $\mathbb{R}$-linear space spanned by the k-fold exterior products

$$v_1 \wedge ... \wedge v_k, \quad v_i \in \mathbb{R}^n$$

which we call the pure or decomposable k-vectors. The exterior product is multilinear and has the alternating property

$$(...v_i \wedge .. \wedge v_j..) = -(...v_j \wedge ... \wedge v_i...)$$

or equivalently

$$...v_i \wedge ... \wedge v_i... = 0.$$

Given a basis $e_1, ..., e_n$ of $\mathbb{R}^n$ a basis of $\Lambda^k(\mathbb{R}^n)$ is given by the $\binom{n}{k}$ elements

$$e_{i_1 i_2 ... i_k} := e_{i_1} \wedge e_{i_2} \wedge ... \wedge e_{i_k}$$

with $i_1 < i_2 < .. < i_k$.

Pure k-vectors correspond to k-dimensional subspaces of $\mathbb{R}^n$. To be precise, let $v_1, ..., v_k$ be linearly independent then the kernel of the linear map $\mathbb{R}^n \to \Lambda^{k+1}(\mathbb{R}^n)$

$$v \mapsto v \wedge (v_1 \wedge ... \wedge v_k)$$

can be shown to be the linear span of $v_1, ..., v_k$. On the other hand form multilinearity follows that if $w_1, ..., w_k$ is another basis of $\text{span}(v_1, ..., v_k)$ then

$$w_1 \wedge ... \wedge w_k = \det T v_1 \wedge ... \wedge v_k$$

where T is the base change map.

In the end there is thus an injective map

$$\gamma : Gr(k, n) \mapsto \mathbb{P}(\Lambda^k(\mathbb{R}^n))$$

form the set $Gr(k, n)$ of k-dimensional subspaces of $\mathbb{R}^n$ and the projective space

$$\mathbb{P}(\Lambda^k(\mathbb{R}^n)) = \Lambda^k(\mathbb{R}^n)/ \sim$$

where $\sim$ is the equivalence relation identifying two pure k-vectors that are nonzero multiples of each other

$$v_1 \wedge ... \wedge v_k \sim w_1 \wedge ... \wedge w_k \Leftrightarrow v_1 \wedge ... \wedge v_k = C\, w_1 \wedge ... \wedge w_k, \quad C \neq 0.$$

γ is called the *Plücker embedding.*

3.2 Inner product

On $\Lambda^k(\mathbb{R}^n)$ one may introduce an inner product which on the pure k-vectors is given by the Gram determinant

$$\langle v_1 \wedge ... \wedge v_k, w_1 \wedge ... \wedge w_k \rangle := \det(\langle v_i, w_j \rangle) \tag{5}$$

and is then linearly extended to all k-vectors.

Now that we have an inner product we can as usual define angles between k-vectors as the arccosine of normalized inner product. Especially for elements of the Grassmannian, we can define such angles. Namely let L_1 and L_2 be two k-dimensional subspaces with bases $v_1, ..., v_k$ and $w_1, ..., w_k$ then

$$\angle(L_1, L_2) = \arccos\Big(\frac{\langle v_1 \wedge ... \wedge v_k, w_1 \wedge ... \wedge w_k \rangle}{\sqrt{\langle v_1 \wedge ... \wedge v_k, v_1 \wedge ... \wedge v_k \rangle \langle w_1 \wedge ... \wedge w_k, w_1 \wedge ... \wedge w_k \rangle}}\Big). \tag{6}$$

One may convince oneself, that of course this number does not depend on the chosen basis.

4 Grassmannian and partial correlation

Let as before $\xi^1, ..., \xi^n$ be sample vectors with partial correlation coefficient given by formula (3)

$$r_{12;3..m} = -\frac{R_{12}}{\sqrt{R_{11}R_{22}}}.$$

Further define the linear hulls $L_1 = L(\xi^2, ..., \xi^k), L_2 = L(\xi^1, \xi^3, ..., \xi^k)$. Since $r_{ij} = \langle \xi^i, \xi^j \rangle$ we immediately discover that

$$-R_{12} = \langle \xi^2 \wedge ... \wedge \xi^k, \xi^1 \wedge \xi^3 \wedge ... \wedge \xi^k \rangle$$

$$R_{11} = \langle \xi^2 \wedge ... \wedge \xi^k, \xi^2 \wedge ... \wedge \xi^k \rangle$$

$$R_{22} = \langle \xi^1 \wedge \xi^3 \wedge ... \wedge \xi^k, \xi^1 \wedge \xi^3 \wedge ... \wedge \xi^k \rangle$$

and therefore

$$r_{12;3..m} = \cos(\angle(L_1, L_2)). \tag{7}$$

Eq. (7) is the desired result, which now gives a simple geometric interpretation of partial correlation also for more than one external variable.

Finally we should mention, that the above defined angle between flats can also be computed as the product of the cosine of the principle angles between these flats. For this define sample matrices $\Xi_1 = (\xi^2, ..., \xi^k), \Xi_2 = (\xi^1, \xi^3, ..., \xi^k)$, so that $\Xi_i^T \Xi_j = R_{ij}$.

Then it is known, that the eigenvalues of $R_{11}^{-1} R_{12} R_{22}^{-1} R_{21}$ are $\cos^2(\theta_l)$ where $0 \leq \theta_l \leq \pi/2$ are the principal angles between L_1 and L_2, see e.g. [3], chapter 10.

Therefore

$$\det(R_{11}^{-1} R_{12} R_{22}^{-1} R_{21}) = \prod_{l=1}^{k} \cos^2(\theta_l).$$

On the other hand the left hand side of this equation is just $r_{12;3..m}^2$ and therefore

$$|r_{12;3..m}| = \prod_{l=1}^{k} \cos(\theta_l), \tag{8}$$

which is also known as the Fubini-Study distance between L_1 and L_2.

References

[1] BILIN ZENG, Kang Chen; WANG, Cong. Geometric views of partial correlation coefficient in regression analysis. International Journal of Statistics and Probability, 2017, 6. Jg., Nr. 3.

[2] DEMESHEV, Boris; GNILOVA, Olya. How Gauss and Markov Met Pythagoras: Geometry in Econometrics. 2018.

[3] EATON, Morris L., et al. Chapter I0: Canonical correlation coefficients. In: Multivariate Statistics. Institute of Mathematical Statistics, 2007. S. 403-463.

[4] KOZLOV, Sergey Emel'yanovich. Geometry of the real Grassmannian manifolds. Parts I, II. Zapiski Nauchnykh Seminarov POMI, 1997, 246. Jg., S. 84-129

[5] LANGSRUD, Øyvind. The geometrical interpretation of statistical tests in multivariate linear regression. Statistical Papers, 2004, 45. Jg., Nr. 1, S. 111-122.

[6] NICKODEM, Kyle, et al. Using Geometry to Visualize Abstract Aspects of Statistical Formulae Relevant to Correlation and Regression.

[7] RAHMAN, Najeeb Abdur. A Course in Theoretical Statistics, Charles Griffin and Company, 1968

[8] SAVILLE, David J.; WOOD, Graham R. Statistical methods: The geometric approach. Springer Science and Business Media, 2012.

[9] STUART, Alan; ORD, John Keith; ARNOLD, Steven F. Kendall's Advanced Theory of Statistics: Classical Inference and the Linear Model. Volume 2A. John Wiley, 2004.

[10] THOMAS, Guy; O'QUIGLEY, John. A geometric interpretation of partial correlation using spherical triangles. The American Statistician, 1993, 47. Jg., Nr. 1, S. 30-32.

Bisher erschienene Research Papers:

Nr. 1 | Februar 2017
Berlemann, Michael; Christmann, Robin.
The Role of Precedents on Court Delay.
Evidence from a Civil Law Country.

Nr. 2 | September 2017
Matthes, Roland.
A note on the Saito-Kurokawa lift for Hermitian forms.

Nr. 3 | Februar 2018
Christmann, Robin.
Prosecution and Conviction under Hindsight Bias in
Adversary Legal Systems.

Nr. 4 | März 2019
Broere, Mark; Christmann, Robin.
Takeovers, Shareholder Litigation, and the Free-riding Problem

Nr. 5 | März 2019
Matthes, Roland.
Vector-valued Cusp Forms and Orthogonal Modular Forms

Nr. 6 | März 2020
Matthes, Roland.
A Geometric View on Linear Regression and Correlation Tests

Nr. 7 | März 2020
Matthes, Roland.
A Note on the Geometry of Partial Correlation and the Grassmannian

Leibniz FH

Leibniz-Fachhochschule
Expo Plaza 11
30539 Hannover

leibniz-fh.de