Characterisation and Identification of multivariate latent manifolds: Analytically resolving factor score indeterminacy

Landon Hurley[1]

July 3, 2023

[1] ljrhurley@gmail.com

Linear Factor Analysis I

- Statistical estimation is a problem of adequate summarisation upon stable regular structures.

- Regularity of structure enables generalisability: the structure of the sample is a homogeneous representation of a larger population.

- Linear models are defined upon the Frobenius norm-space: ℓ_2. This is due to mathematical origins in function approximation.

- Constructively, this results in linear models possessing Gaussian errors: combined with sampling assumptions, we obtain an a.s. convergent affine-linear Hilbert space.

Linear Factor Analysis I

- This is also a problem though: if we observe non-linear data scores, the Gaussian error assumption is false, compromising the Gauss-Markov theorem.

- This results in biased estimators which are not stable and do not possess minimum variance, or tight, estimation bounds.

- The generalised linear model developed to address in linear models non-linear scores, by assuming a bilinearity condition (the central limit theorem).

- Under the CLT, exponential distributions converge to normality under the weak lln. This allows us to identify both the ranking and scoring of individual elements uniquely (as a Hilbert space; Riesz representation theorem)

Roadmap of Talk

Behind the topology of Linear Factor Analysis as an estimation problem

Why must Latent Variables be Gaussian?

Estimating non-Gaussian latent variables with linear models

Topology of the Riemannian manifold to prove a consistent Hadamard solution to Factor Score Indeterminacy

Extensions to address Factor Rotation indeterminacy

Linear topology upon a population I

Table: Presentation of the observation of either a linear or non-linear manifold wrt the ranks and scores of upon a collection of random variables and their errors.

	Ranks	Ranks
Scores	Linear	Non-linear
Linear	(1)	(2)
Non-linear	(3)	(4)

Discussion outline I

- By Nelder and Wedderburn (1972), we must explicitly identify the unique monotonically non-decreasing link function to ensure stability.

- If mis-chosen, $g(\cdot)$ defines a convergent yet biased estimator. By CLT, we obtain asymptotically unbiased but locally biased results. By IRLS of Expected Fisher Information, $\mathcal{I}^{\frac{-1}{2}}$ matrix, the results are strictly false if the estimator is biased though.

- Thus, finite sample differences become gratuitous, and we would expect such findings to fail to generalise (easy instantiation of the failed application of the Berry-Essen Theorem for Sum Scores).

- Bilinearity upon rank and score, is a population condition upon the ℓ_2-norm space. A unique estimator of the most accurate rankings only is observed when the score function is known exactly: $\lim_{n\rightarrow\infty^+} F_X(x) = r, F_X^{-1}(r) = x$.

Topology of our problem I

- We often cannot uniquely identify adequate regularity in F: trivial examples are Olsson (1979) MLE, non-parametric estimators, as well as Heywood cases. The lack of analytically expressible second order regularity for Owen (2001) is also a problem (bootstrapping).

- We lack a corresponding complete Hilbert function space for the permutations upon a sample. The empirical approach attempts to correct this, but has no regularity outside the sample itself. This is also why Mokken and Lewis (1982) is unidimensional: ties upon linear mappings.

- Part of the problem lies in the lack of a suitable probabilistic mapping for rankings: $S_n = n!$ is insufficient for handling multiple covariates, as the linear combination of elements typically results in ties.

Topology of our problem II

- Grice (2001) pointed our a parable about multiple score distributions fit equally well (maximise ℓ_2 fitness) but possess divergent rankings.

- The asymptotic linearity of the ranking though leaves us unable to freely estimate ranking upon finite samples. We need a means of evaluating finite samples (n) upon finite tests (p).

Roadmap of Talk

Behind the topology of Linear Factor Analysis as an estimation problem

Why must Latent Variables be Gaussian?

Estimating non-Gaussian latent variables with linear models

Topology of the Riemannian manifold to prove a consistent Hadamard solution to Factor Score Indeterminacy

Extensions to address Factor Rotation indeterminacy

Gaussian latent variables I

- The Gaussian nature of errors in function approximation is a consequence of Eckart and Young (1936) being constructed as a Hadamard estimation problem.

- We construct p estimating equations for each of q latent dimensions, averaging over n sample individuals.

- By the central limit theorem, even if p variables are not normally distributed, their asymptotic weighted linear combination is still Gaussian. Thus, all latent variables are asymptotically Gaussian.

- As linear functions with individual p marginal Gaussian errors, we expect the data structure to be unbiased and possess minimum variance, and the corresponding latent variables to therefore be Gaussian.

Gaussian latent variables II

- These conditions are, upon the test item population, identical to that of the Rasch (1961) model: we expect Test scores to uniquely determine the ranking of the sample proportionately to their true score $\theta_i^{n \times q}$, $i = 1, 2, \ldots, n$.

- If the observed items do not share common discrimination though, there is no constant ΔT for test score. Once the 1PL is generalised, there is no longer bilinearity.

- As a sub-space embedding by the Whitney theorem, the contraction upon $\mathbf{X}_{n \times p}^{\top} \times \mathbf{X}^{n \times p} \rightarrow \theta^{n \times q}$, we lose the ability to define a unique linear score if finite sample bilinearity is lost.

Gaussian latent variables III

- Wilson (1928) first revealed this, by identifying the loss of uniqueness upon quadratic polynomials and their lack of singular roots $\pm\lambda_j$, and Bochner expressed this as

$$\lim_{p\to\infty^+} \frac{p^2 - p}{2} + p \leq p^2,$$

which only achieves equality in the limit. Thus, factor score indeterminacy.

Stability of Gaussian latent variables I

- A natural next question is then: is this asymptotic normality upon $\theta^{n \times q}$ stable wrt n & p?

- Assuming *i.i.d.*, the population of test items contractively embedded upon $\theta^{n \times p}$, $n \leq \infty^{+}$ is conditionally independent of all other samples, and thus converges via the Glivenko (1933)-Cantelli (1933) theorem in the same procedure. A contractive mapping (Eckart & Young, 1936) ensures the linearity of the observed space and its sub-space.

- Next, consider relaxing to $p < \infty^{+}$. Under this relaxation, the linear embedding of the Gaussian errors upon $\mathbf{X}^{n \times p} \mapsto \theta^{n \times q} + \epsilon, \epsilon \sim \mathcal{N}(\mu^{q \times 1}, \Sigma^{q \times q})$, the observed values follows through upon the Gaussian latent variables.

- Thus, if the latent variables are normally distributed, we expect to obtain unbiased tight results upon finite samples wrt both $n \& p$. This supports the claim that the GOF χ^2-test should hold, uniformly via Glivenko-Cantelli being the most powerful misspecification detector.

Non-Gaussian latent manifolds I

- However, the distribution of the latent manifold is an assumption, not a guarantee unless performed upon the population wrt the test items. This limit is therefore not stable: asymptotically normal latent variables are not expected to be noramlly distributed upon finite samples when defined by only a few indicators.

- We offer a resolution to this problem by identifying a linear manifold and corresponding probability structure which is orthonormal to the ℓ_2 definition of the problem.

- Explicitly, we address the problem of Hadamard estimation problem as the ML or linear Gauss-Markov (GM) estimate of the latent scores which are ranked most accurately.

- This process is achieved by separating and defining a linear topology upon the ranks, and noting that, under certain general conditions, even non-linear scores are almost surely linearly ranked.

Non-Gaussian latent manifolds II

- The first problem consists of defining a sufficient domain for the problem: S_n is typical, but the lack of continuous random variables and their linear combination results in strong probability of ties occurring.

- We require a means of defining a metric space over permutations with ties: $\mathcal{M}_n = \{n^n - n\}_m$.

- Note that we explicitly exclude the now possible n permutation events which are degenerate from the population.

- Kemeny (1959) first identified a metric space for this, however his expression was too similar to Kendall (1938), and failed to recognise that the Kemeny space was a Hilbert space.

- Upon a Hilbert space, we achieve almost sure convergence via the Borel-Cantelli lemma, with a complete and unique probabilistic mapping for all n.

Kemeny metric I

$$\tau_\kappa(x,y) = -\tfrac{2}{n^2-n}\sum_{k=1}^{n}\sum_{l=1}^{n}\kappa(x)_{kl}\odot\kappa^{\mathsf{T}}(y)_{kl} = -\tfrac{2}{n^2-n}\left(\rho_\kappa(x,y)-\tfrac{n^2-n}{2}\right), \{x,y\}\in\overline{\mathbb{R}}^{n\times 1}. \tag{1a}$$

$$\rho_\kappa(x,y) = \frac{n^2-n}{2} + \sum_{k,l=1}^{n}\kappa_{kl}(x)\odot\kappa_{kl}^{\mathsf{T}}(y),\ \ k,l = 1,\ldots,n. \tag{1b}$$

$$\kappa_{kl}(x) = \begin{cases}\sqrt{.5} & \text{if } x_k > x_l\\ 0 & \text{if } x_k = x_l,\\ -\sqrt{.5} & \text{if } x_k < x_l\end{cases} \tag{1c}$$

$$\kappa_{kl}(y) = \begin{cases}\sqrt{.5} & \text{if } y_k > y_l\\ 0 & \text{if } y_k = y_l,\\ -\sqrt{.5} & \text{if } y_k < y_l\end{cases} \tag{1d}$$

$$\dot{\sigma}_\kappa^2(\mathcal{M}_n) = \frac{(n-1)^2(n+4)(2n-1)}{18n}, \tag{2a}$$

$$\sigma_\kappa^2(X) = \tfrac{2}{n(n-1)}\left(\sum_{k=1}^{n}\sum_{l=1}^{n}\kappa_{kl}(X)\kappa_{kl}^{\mathsf{T}}(X)\right) \equiv \tfrac{2}{n(n-1)}\sum_{l,k=1}^{n}\kappa_{kl}^2(X). \tag{2b}$$

- The Kemeny metric is a Hilbert space, and also a Gauss-Markov estimator. It is affine-linear over monotonically non-decreasing functions as well.

- It operates upon $\mathcal{M}_n$, the set of all permutations upon the extended real line, and converges for all n almost surely.

- $S_n \subset \mathcal{M}_n$ as well. It addresses the problem of the probability of observing ties uniquely.

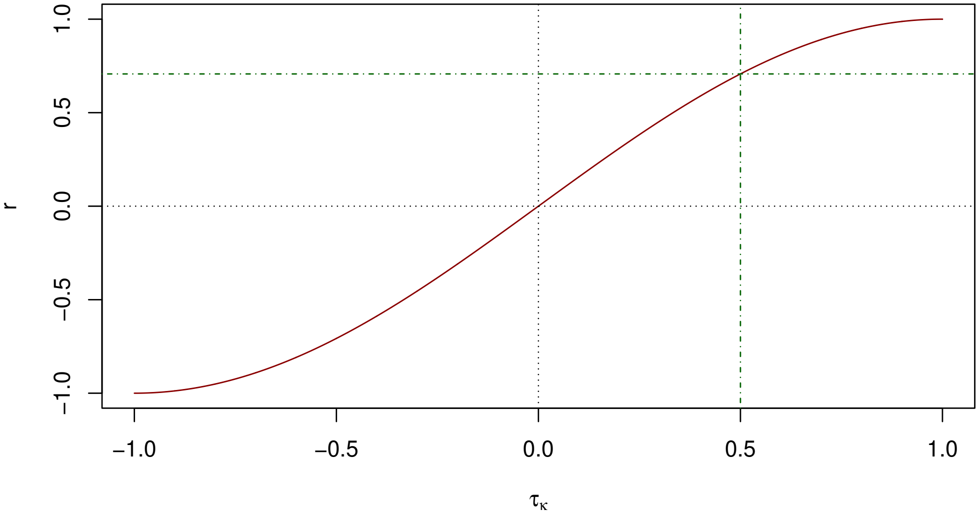
Kemeny metric II

- This allows us to express a distance matrix between all vectors $X^{n \times p}$ upon the sample, and an affine-linear transformation of the distances provides a correlation measure τ_κ as well.

- In Kendall (1948, p. 129) the following equivalence was established:

$$r_{X,Y} = \sin\left(\tau_b(X,Y) \cdot \frac{\pi}{2}\right), \tag{3}$$

Kemeny metric III



Important to note:

1. $\ell_\kappa \neq \ell_2$, but can be isometrically embedded, as all Hilbert spaces are equivalent to the Euclidean distance.
2. $U(\mathcal{M}_n)$ follows a Beta-Binomial distribution which is only asymptotically normally distributed: stably strictly sub-Gaussian
3. It defines a linear function space which can be approximated by ℓ_2 though.
4. Define $\Xi^{p\times p}$ as a Gramian covariance matrix of p input vectors which may possess non-linear scores.
5. Together with the affine-linearity over monotone functions, including $g \equiv I(\cdot)$., ℓ_2 expresses ρ_κ satisfactorily: this enables a linear short-mapping.
6. Mapping $\Xi^{p\times p} \mapsto \theta^{n\times q}$ is valid, but it does not necessarily follow that the embedding is Gaussian. For example $\Phi^{q\times q}$ becomes the inner-product of the linear embedding of strictly sub-Gaussian latent variables, which upon $\ell_2 \approx \Xi$ denotes a second correlation, $\rho_\kappa \sim \Phi^{q\times q}$.

Roadmap of Talk

Behind the topology of Linear Factor Analysis as an estimation problem

Why must Latent Variables be Gaussian?

Estimating non-Gaussian latent variables with linear models

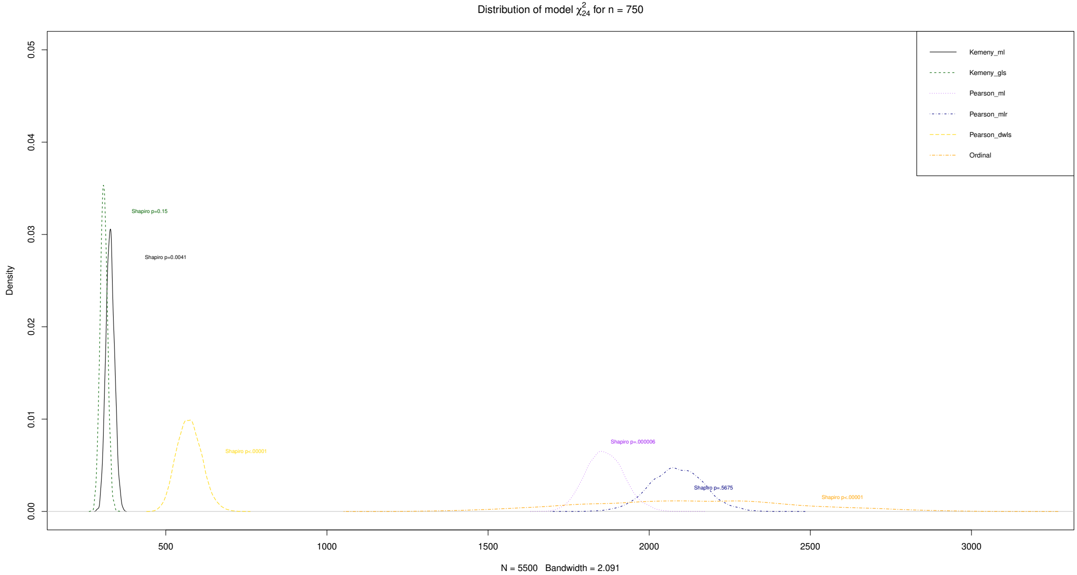
Topology of the Riemannian manifold to prove a consistent Hadamard solution to Factor Score Indeterminacy

Extensions to address Factor Rotation indeterminacy

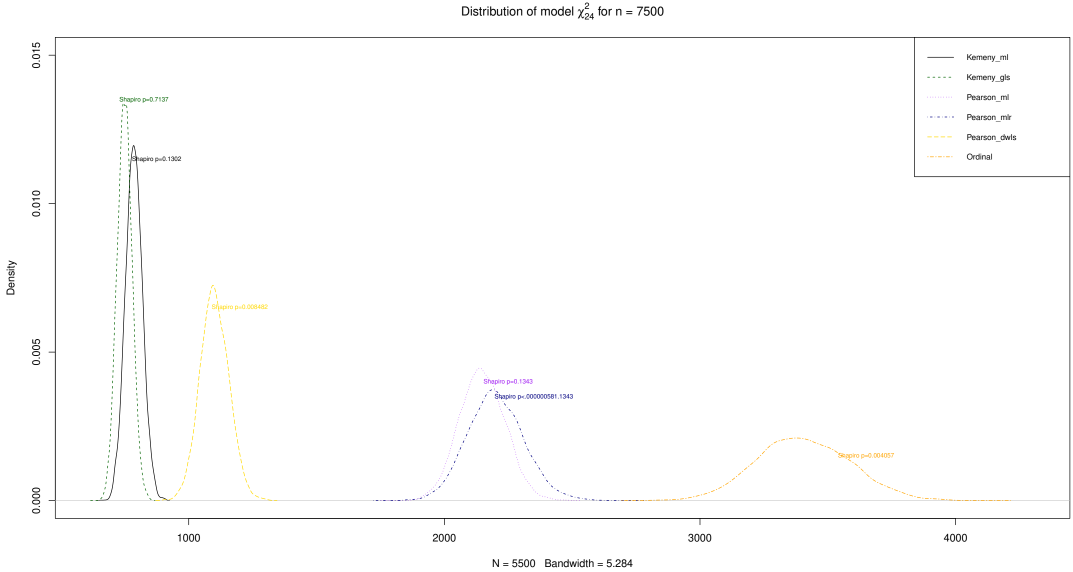
Problem as currently stands I

- Does resolve the possibility of non-linear score embeddings upon linear rankings.
- Does NOT resolve identification of latent variable scores, other than showing they may not be Gaussian.
- So what is the relationship upon the $\frac{p^2-p}{2} + p \ll p^2$ known parameters of Ξ and the $\frac{p^2-p}{2} + p \ll p^2$ known parameters of Σ?
- First, note that $\frac{p^2-p}{2} + p + \frac{p^2-p}{2} = p^2$
- This leaves the estimation problem just-identified, but as the embedding cannot be a contraction, how can the two orthonormal metric spaces be linearly optimised upon?

Supporting our claim: PI



Supporting our claim: PII



Gaussian latent variables I

Table: Distributions of the likelihood-ratio tests for the Holzinger data set under different estimators, Ordinal defining the combination of polychoric and DWLS estimators, for various *n*.

n		mean	sd	median	mad	min	max	range	skew	kurtosis
n = 301	Kem_ml	135.78	8.54	135.41	8.50	107.94	172.14	64.20	0.19	0.06
	Kem_gls	125.25	7.10	125.11	7.07	100.22	156.27	56.05	0.02	0.05
	Pear_ml	762.03	39.11	759.76	38.54	642.64	952.66	310.03	0.36	0.28
	Pear_gls	838.45	52.84	837.61	53.23	654.54	1037.89	383.35	0.09	0.05
	Ken_ml	238.11	25.11	236.03	24.41	165.14	355.81	190.67	0.51	0.52
	Ken_gls	836.07	165.33	837.49	179.24	388.66	1371.87	983.21	-0.02	-0.56
n = 750	Kem_ml	327.49	13.01	327.15	13.05	284.63	371.73	87.10	0.11	-0.09
	Kem_gls	307.78	11.15	307.55	11.11	267.35	354.09	86.75	0.09	-0.01
	Pear_ml	1862.46	59.63	1861.04	60.92	1660.84	2146.83	485.99	0.16	0.08
	Pear_gls	2086.23	83.36	2086.02	83.58	1734.29	2443.53	709.23	-0.01	0.11
	Ken_ml	573.25	38.94	572.07	39.50	458.98	743.20	284.21	0.26	0.06
	Ken_gls	2128.23	326.40	2126.12	337.12	1209.75	3113.24	1903.49	0.06	-0.40
n = 7500	Kem_ml	3209.740	41.008	3209.908	40.780	3055.841	3345.246	289.405	-0.007	-0.056
	Kem_gls	3052.594	36.288	3052.480	36.659	2923.028	3175.430	252.402	0.013	-0.090
	Pear_ml	18402.979	186.499	18403.916	189.628	17784.339	19041.252	1256.913	0.064	-0.125
	Pear_mlr	20865.061	266.452	20865.209	274.161	19974.056	21818.144	1844.088	-0.021	-0.118
	Pear_dwls	5610.712	120.587	5607.139	122.131	5218.355	6090.508	872.153	0.089	0.006
	Ordinal	21664.151	1192.641	21666.358	1181.759	16652.545	27453.147	10800.602	0.012	0.483

Intuitively defining our problem I

- The first solution is to characterise $\hat{\theta}$ as to obtain a set of scores which are ordered most accurately upon finite samples of indicators.

- To do this requires recognising the Riemannian manifold upon the domain of the estimator: linearly separate scores and ranks which are independently estimated, but share a common unique $\hat{\theta}^{n \times q} \rightarrow \theta^{n \times q}$.

- Asymptotically wrt p, said latent variables are normally distributed, but otherwise strictly sub-Gaussian: we seek the LS solution which satisfies the GM ranking of $\hat{\theta}^{n \times q}$.

- The KKT score solution is actually Ridge regression, which is generally accepted to be non-uniquely identified.

Roadmap of Talk

Behind the topology of Linear Factor Analysis as an estimation problem

Why must Latent Variables be Gaussian?

Estimating non-Gaussian latent variables with linear models

Topology of the Riemannian manifold to prove a consistent Hadamard solution to Factor Score Indeterminacy

Extensions to address Factor Rotation indeterminacy

KKT estimation of the latent factors I

Definition

An objective function $f : \overline{\mathbb{R}}^p \to \mathbb{R}$ and the constraint functions $g_i : \mathbb{R} \to \mathbb{R}$ and $h_j : \mathbb{R}^n \to \mathbb{R}$ possess sub-derivatives at a point $x^* \in \mathbb{R}^n$. Should x^* be a local extrema while simultaneously satisfying regularity conditions, there exists constants $\mu_i (i = 1, \ldots, m)$ and γ_j^2, which are the Karush-Kuhn Tucker multipliers, whose solution must satisfy the following four conditions:

1. Stationarity
2. Primal feasibility
3. Dual feasibility
4. Complementary slackness.

KKT estimation of the latent factors II

- The theorem of the Karush-Kuhn-Tucker conditions holds as follows:

Theorem

If there exists a solution x^* to the primal problem, and a solution (u^*, v^*) to the dual problem, which together satisfies the Karush-Kuhn-Tucker conditions (Definition 1), then the problem pair has strong duality, and $x^*, (u^*, v^*)$ is a solution pair to the prime and dual problems.

- Note that strong duality is the population bilinearity condition under the CLT: the gap between the score solution and rank solution is minimised, and must converge to 0 upon the population as a function of $p \to \infty^+$.

Strong Duality I

Theorem

A strong duality between solutions to the minimisation of the Kemeny norm (primal problem) and the minimisation of the Frobenius norm (dual problem) defines a duality gap equal to 0, and thus establishes the projective geometric duality between the two topologies to be a strongly duality.

- In Kendall (1948, p. 129) the following equivalence was established:

$$r_{X,Y} = \sin\left(\tau_b(X,Y)\cdot\frac{\pi}{2}\right), \tag{4}$$

- However, strictly speaking, if $\tau_b \subset \tau_\kappa$, then the left-expression is only true in the limit:
- For finite samples (or $X, Y \in \overline{\mathbb{R}}^{n\times 2}$,) estimator is a linear combination of n elements' ranks, not scores.

Strong Duality II

- This is not a problem, as $\lim_{n \to \infty^+} \rho_\kappa(X, Y) \to r(X, Y)$ by the CLT.

- By Brouwer's fixed point theorem, collinearity upon the similarity measure occurs at three fixed mapping points $\{-1, 0, 1\}$, the outer two of which are excluded (as they denote collinear solution points) from $\mathcal{M}_n$.

- Thus, there remains one fixed point affine-linearly invariant solution point common to both spaces, which denotes complete independence relative to affine linear transformations. This is the origin which is coincidental upon each orthonormal abelian function space.

- Empirically, it reflects the convergence of the mediand the mean.

Strong Duality III

- The duality gap is then the difference given by

$$\inf_{x \in X} [F(x, 0)] - \sup_{y^* \in Y^*} [-F^*(0, y^*)], \tag{5}$$

  where F^* is the convex conjugate in both variables. Note that per equation 4 is obtained $d^* = r_{X,Y}$ and $p^* = \sin\left(\tau_{\kappa|X,Y} \cdot \frac{\pi}{2}\right)$, such that $p^* - d^* = 0$, and thus strong duality is established for the solution to both orthonormal linear topologies.

- Consider d^* to contain not only unique parametric but also Tikhinov-regularised solution spaces, then there exists multiple solutions for the infinite set $\gamma_M^2 = \{\gamma_j^2\}_{j=1}^p$ which minimise f subject to the satisfaction of the isocontour of the Euclidean distance.

Strong Duality IV

- However, there are two orthonormal criterion upon α, the regularity parameter space, and thus a unique solution is obtained at the point x^* which minimises both the Euclidean and Kemeny metric spaces of the projection and the target, defined at $f \mid \inf g^*$. Thus is defined a manifold sub-space of all solutions which together denote the duality gap, by constraint equality of 0 upon the population, and is otherwise a unique infinimum.

- Weak and strong duality thus holds as follows: there exists a unique Kemeny distance exists between all finite vectors of length n.

- The solution to the Tikhinov-regularised duality obtains an infinite set of solution points (p coefficient parameters) which all produce an identical finite distance upon ℓ_2. However, upon this image of the optimal finite isocontour exists a unique solution which also minimises the Kemeny distance.

Strong Duality V

- By Farkas' Lemma, there is always one of the following alternatives, and the only non-imaginary root is a real unique solution. The alternative Fredholm alternative results in the imaginary solutions in Quantum Mechanics.

Non-Gaussian latent variables I

- Solving the KKT in this fashion is sufficient, but inefficient: This is very similar to the 2-step Polychoric correlation as well

- We can use this to address a number of estimation problems while satisfying the Gauss-Markov expectations wrt the solution space.

- Of course, this is leveraging Slater's condition: thus it remains fixed as two separate estimation problems.

- We need a way to resolve this as a one step problem: we ended up resolving this as a Weighted Least Squares problem: Browne (1984)

- Aitken (1936) first introduced WLS, however it resolves upon the asymptotically obtainable (and thus the Feasibility of GLS) of the true covariance matrix.

Non-Gaussian latent variables II

- This produces consistent, but not efficient solutions. In the terms of KKT though, we possess two measures of covariance which must be solvable for finite samples upon $\hat{\theta}$: thus, we obtain a population of dispersions which possess the scores which are ordered most accurately upon the sample.

- This is why we compare ML and GLS decompositions: asymptotically, the order must be minimsed in order to identify the correct scoring. However, no scoring may be obtained which does not also minimse the Kemeny distance of error from the target.

- Our problem is that the Kemeny metric does not allow direct interpretation of the projection regression function upon the original domain.

MLE upon an unusual topological space I

- An alternative function is offered by equation 4 though. Upon equation 1c exists $\kappa(X)^{n \times n}$: if we sum over the rows and transpose, we obtain a vector $X_*^{n \times 1}$ constructed upon the ranks, rather than the scores.

- This provides the connection for ρ_κ as the ℓ_2 approximation of Ξ by LS.

- Further, it allows us to define the target orthonormally upon the same function domain in terms of n individual elements:

- This is $\omega_i^{n \times 1} = \mathtt{diag}(\Omega^{n \times n}), i = 1, \ldots, n$, where the relative differences are produced by the errors in rank orderings of $\omega_i \propto 1 + \frac{(\hat{Y}_*^{i \times 1} - Y_*^{i \times 1})^2}{(n^2 - n)^2}$.

$$\mathbf{V}^{p \times p} = \mathbf{X}_{n \times p}^{\intercal} \Omega_{n \times n} \mathbf{X}_{n \times p} \tag{6}$$

$$\mathbf{V} = \Psi + \Lambda \Phi^{\intercal} \Lambda. \tag{7}$$

MLE upon an unusual topological space II

- This expresses the errors in the expected rankings of the predicted scores Y, and iteratively reweights the scores until the median expectation of error (the Kemeny distance) is minimised and stable. Thus, the mean prediction converges to the unbiased estimate of the median.

- Formally, this approach satisfies the KKT saddlepoint solution as a single procedure, and provides an estimator whose solution is a Gauss-Markov linear solution.

- A linear GM solution is, by definition, a maximum likelihood estimator: thus, by minimising the errors in ordering with the most accurate greedy score approximations, we obtain a unique solution equivalent to Farkas' lemma, almost surely.

Factor Score Determinations I

- Thus by either the KKT or the MLE, we obtain a unique solution upon the Kemeny objective function to the question: what are the unique best estimated latent factor scores which are most correctly ordered?

$$\mathbf{w}_{\kappa}^{p\times q} = (\Xi)_{p\times p}^{-1} \Lambda_{\kappa}^{p\times q} \Phi_{\kappa}^{q\times q} \tag{8a}$$

$$\theta_{*}^{n\times q} = \mathbf{X}^{n\times p}\mathbf{w}^{p\times q} + \epsilon, s.t.\ E(\epsilon_{*}) = 0, \tag{8b}$$

- If we obtain $\hat{\theta}^{n\times q}$, the GM solution to the best ordering of scores, we next require a set of linear embeddings which best approximate these scores. Solving equation 8b upon S and $\hat{\Lambda}^{p\times q}$ produces $\hat{\theta}^{n\times q}$ such that the embeddings produce the unique smallest distance to $\min_{\rho_{\kappa}}(\hat{\theta}_{*}^{n\times q}, \hat{\theta}^{n\times q})$, we obtain a unique solution via the KKT solution conditions.

- By Farkas' lemma, such a solution always exists as well. Consider a set of factor score estimates: the one which maximises τ_{κ} between the test scores and the latent scores also satisfies the KKT conditions.

Roadmap of Talk

Behind the topology of Linear Factor Analysis as an estimation problem

Why must Latent Variables be Gaussian?

Estimating non-Gaussian latent variables with linear models

Topology of the Riemannian manifold to prove a consistent Hadamard solution to Factor Score Indeterminacy

Extensions to address Factor Rotation indeterminacy

Factor rotation indeterminacy I

- Require a means of assessment or estimation of the $\mathrm{T}^{q\times q}$ rotation elements from the $\Lambda_{p\times q}$ initial factor loading matrix, where p is the number of observed variables and q is the estimated number of latent factors.

- We only provide the identification of a unique solution to the free parameter for the oblimin rotation family ζ (Harman, 1960; Jennrich & Sampson, 1966; Jennrich, 1979).

$\Sigma^{p\times p}$ the Pearson variance-covariance matrix for p random variables,

$\Xi^{p\times p}$ the Kemeny variance-covariance matrix for p random variables

$\Lambda_{p\times q}$ the latent factor loadings estimated upon the Kemeny matrix,

$\breve{\Lambda}_{p\times q}$ the latent factor loadings estimated upon the Pearson correlation matrix, under the Kendall sinusoidal transformation.

$\lambda_{a,b}$ elements $\lambda_{a,b} \in \Lambda_{p\times q} \forall \{a, b\}_{a=1,b=1}^{p,q}$

$\breve{\lambda}_{a,b}$ elements $\breve{\lambda}_{a,b} \in \breve{\Lambda}_{p\times q} \forall \{a, b\}_{a=1,b=1}^{p,q}$

Factor rotation indeterminacy II

$$\begin{aligned}
0 &= \underbrace{\sum_{a=1}^{p}\sum_{b=1}^{q}\lambda_{ab}^{4}}_{f} - \frac{\zeta}{p}\underbrace{\sum_{b=1}^{q}\left(\sum_{a=1}^{p}\lambda_{ab}^{2}\right)^{2}}_{g} \\
0 &= f - \frac{\zeta}{p}\cdot g
\end{aligned} \tag{9a}$$

$$\begin{aligned}
0 &= \underbrace{\sum_{a=1}^{p}\sum_{b=1}^{q}\breve{\lambda}_{ab}^{4}}_{f^{*}} - \frac{\zeta}{p}\underbrace{\sum_{b=1}^{q}\left(\sum_{a=1}^{p}\breve{\lambda}_{ab}^{2}\right)^{2}}_{g^{*}} \\
0 &= f^{*} - \frac{\zeta}{p}\cdot g^{*}
\end{aligned} \tag{9b}$$

Factor rotation indeterminacy III

$$0 = \frac{\zeta}{p}\Big(\frac{p}{\zeta}\sum_{a=1}^{p}\sum_{b=1}^{q}\lambda_{a,b}^{4} - \sum_{b=1}^{q}(\sum_{a=1}^{p}\lambda_{a,b}^{2})^{2}\Big) - \frac{\zeta}{p}\Big(\frac{p}{\zeta}\sum_{a=1}^{p}\sum_{b=1}^{q}\breve{\lambda}_{a,b}^{4} - \sum_{b=1}^{q}(\sum_{a=1}^{p}\lambda_{a,b}^{2})^{2}\Big)$$

$$\zeta = p\Big(\frac{f-f^{*}}{g-g^{*}}\Big)$$

$$\zeta = p\left(\frac{\sum_{a=1}^{p}\sum_{b=1}^{q}\lambda_{a,b}^{4} - \sum_{a=1}^{p}\sum_{b=1}^{q}\breve{\lambda}_{ab}^{4}}{\left(\sum_{b=1}^{q}\lambda_{ab}^{2}\right)^{2} - \sum_{b=1}^{q}\left(\sum_{a=1}^{p}\breve{\lambda}_{ab}^{2}\right)^{2}}\right). \tag{9c}$$

- Thus follows a unique solution upon the estimated and almost surely observed un-rotated factor loadings constructed from the Kemeny and Euclidean similarity matrices, each of which are presumed observed upon the necessary unbiased minimum variance assumptions upon a p-dimensional multivariate observed manifold.

Factor rotation indeterminacy IV

- There is a clear paradox though, in that a unique ζ for the entire system of linear equations would thus define a single common correlation coefficient for all latent variable pairs, which is clearly nonsensical.

- We must therefore consider the pairwise evaluation of each element $\phi_{r,s}$ in the latent variable correlation matrix, using a distinct $\zeta_{r,s}$ using the above definition upon the individual paired subset of latent variables relative to all. We note that this procedure explicitly resolves the iterative non-linear resolution which must otherwise be conventionally employed to solve for T (Jennrich, 2002).

- Assume labels $(r, s) \in \binom{q}{2}$, denoting the column-wise paired vectors of length q which represent the factor loadings of each latent variable upon a given correlation between the respective latent factors (thus $\phi_{r,s} \in \Phi_{r,s},\ \ r, s \in \binom{q}{2}$), and each of the $\frac{n^2-n}{2}$ elements may be solved for (Jennrich & Sampson, 1966, p. 316-317):

Factor rotation indeterminacy V

$$
\begin{aligned}
\zeta_{r,s}^2 &= 1 + 2\phi_{r,s}\delta + \delta^2 \\
\zeta_{r,s} &= \frac{1}{t_r} \\
\delta &= t_s\zeta \\
\phi_{r,s} &= \frac{\zeta_{r,s}^2(t_s^2 - 1) + 1}{2\zeta_{r,s}t_s} \\
\zeta_{r,s}^2 &= 1 + 2\big(\frac{\zeta_{r,s}^2(t_s^2 - 1) + 1}{2\zeta_{r,s}t_s}\big) \cdot \frac{t_s}{\zeta_{r,s}} + (t_s\zeta_{r,s})^2 \\
t_s &= \pm\sqrt{\frac{\zeta_{r,s}^2 - 1}{\zeta_{r,s}}},\ \ s.t,\ \ \zeta_{r,s} \notin \{0, \pm 1\}, \\
\phi_{r,s} &= \frac{\zeta_{r,s}^2\left(\pm\sqrt{\frac{\zeta_{r,s}^2-1}{\zeta_{r,s}}}\right)^2 - \zeta_{r,s}^2 + 1}{2\zeta_{r,s} \pm \sqrt{\frac{\zeta_{r,s}^2-1}{\zeta_{r,s}}}}
\end{aligned}
\tag{10}
$$

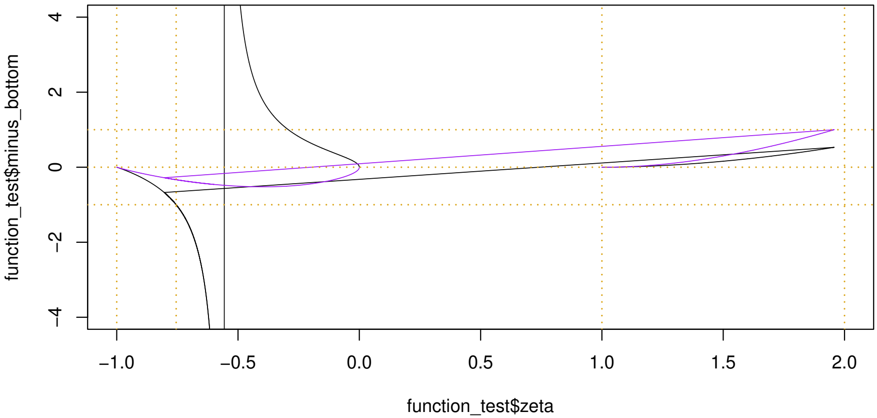
Factor rotation indeterminacy I

- Jennrich (1979) does not fix $\zeta \in [0, 1]$: negative values produced solutions with strong regular convexity. As the estimation of $\hat{\zeta}_{r,s}$ is now an empirical construct conditionally dependent upon the available sample, it follows that solved values of the two orthonormal spaces may produce negative values, especially given the regular linear leptokurtosis of the Kemeny metric space.

- There are 4 equations for each $\phi_{r,s}$ from equation 10 which are plotted in Figure 3, denoting the existence of an over-determined, and thus ill-posed, estimation problem.

Figure: Plot of the domain ζ and correlation coefficient co-image ϕ by equation 10.

Factor rotation indeterminacy II



Factor rotation indeterminacy III

- Resolve this by examining the 4 $\pm$ possible combinations of the quadratic roots.

- Two pairs of the proposed latent roots are algebraically identical (noted by the mixture of the arbitrary signs as either identical, or non-identical), thereby reducing 4 to the 2 with common additions or subtractions. From the remaining set of 2 solutions, there is only one viable image of $\zeta_{r,s}$ which maps onto the interval $[-1, 1]$, as is necessary for a correlation coefficient: thus, one of the two solutions is invalid upon the function $\phi_{r,s}(\zeta_{r,s})$, resolving the over-determined solution set onto a single necessary correlation coefficient. This however reduces the viable space upon which $\zeta_{r,s}$ may exist.

- Presents a unique correlation matrix $\Phi_{q\times q}$ enabling the unique structural pattern conditional upon the observed sample.

Factor rotation indeterminacy IV

- The obtainment of T is of course a direct consequence of the estimated Φ solved upon the estimated Λ to produce a rotated factor loading matrix:

$$\begin{aligned}
\Lambda_{p\times q}\Phi_{q\times q}\Lambda^{\intercal}_{p\times q} &= \breve{\Lambda}_{p\times q}\Lambda^{\intercal}_{p\times q} \\
\Phi_{q\times q} &= (\Lambda^{\intercal}_{p\times q})^{-1}\breve{\Lambda}_{p\times q}\breve{\Lambda}^{\intercal}_{p\times q}(\Lambda_{p\times q})^{-1} \\
\Phi_{q\times q} &= (\Lambda^{\intercal}_{p\times q})^{-1}\Xi_{p\times p}\Lambda^{-1}_{p\times q} \\
\mathrm{T}_{q\times q}\mathrm{T}^{\intercal}_{q\times q} &= (\Lambda^{\intercal}_{p\times q})^{-1}\Xi_{p\times p}\Lambda^{-1}_{p\times q},
\end{aligned}$$

- where T obtained by Cholesky decomposition, wherein is computed the necessary rotation matrix to produce the known target latent correlation matrix Φ as a decomposition of the sufficient similarity matrix Ξ.

Bibliography I

Aitken, A. C. (1936). IV.—on least squares and linear combination of observations. *Proceedings of the Royal Society of Edinburgh*, *55*, 42–48. https://doi.org/10.1017/s0370164600014346

Browne, M. W. (1984). Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, *37*(1), 62–83. https://doi.org/10.1111/j.2044-8317.1984.tb00789.x

Cantelli, F. P. (1933). Sulla determinazione empirica delle leggi di probabilita. *Giorn. Ist. Ital. Attuari*, *4*(421-424).

Eckart, C., & Young, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika*, *1*(3), 211–218. https://doi.org/10.1007/bf02288367

Glivenko, V. (1933). Sulla determinazione empirica delle leggi di probabilita. *Gion. Ist. Ital. Attauri.*, *4*, 92–99.

Grice, J. W. (2001). Computing and evaluating factor scores. *Psychological Methods*, *6*(4), 430–450. https://doi.org/10.1037/1082-989x.6.4.430

Harman, H. H. (1960). *Modern factor analysis*. The University of Chicago Press.

Bibliography II

Jennrich, R. I., & Sampson, P. F. (1966). Rotation for simple loadings. *Psychometrika*, *31*(3), 313–323. https://doi.org/10.1007/bf02289465

Jennrich, R. I. (1979). Admissible values of γ in direct oblimin rotation. *Psychometrika*, *44*(2), 173–177. https://doi.org/10.1007/bf02293969

Jennrich, R. I. (2002). A simple general method for oblique rotation. *Psychometrika*, *67*(1), 7–19. https://doi.org/10.1007/bf02294706

Kemeny, J. G. (1959). Generalized random variables. *Pacific Journal of Mathematics*, *9*(4), 1179–1189. https://doi.org/10.2140/pjm.1959.9.1179

Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika*, 81–93. https://doi.org/10.2307/2332226

Kendall, M. G. (1948). Rank correlation methods..

Mokken, R. J., & Lewis, C. (1982). A nonparametric approach to the analysis of dichotomous item responses. *Applied Psychological Measurement*, *6*. https://doi.org/10.1177/014662168200600404

Bibliography III

Nelder, J. A., & Wedderburn, R. W. M. (1972). Generalized linear models. *Journal of the Royal Statistical Society. Series A (General)*, *135*(3), 370. https://doi.org/10.2307/2344614

Olsson, U. (1979). Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika*, *44*(4), 443–460. https://doi.org/10.1007/bf02296207

Owen, A. B. (2001). *Empirical likelihood*. Boca Raton, FL: CRC Press.

Rasch, G. (1961). Probabilistic models for some intelligence and attainment tests. *Information and Control*, *4*(4), 382. https://doi.org/10.1016/s0019-9958(61)80061-2

Wilson, E. B. (1928). Review: The abilities of man, their nature and measurement. *Science*, *67*(1731), 244–248. https://doi.org/10.1126/science.67.1731.244