Skewness projection pursuit: past, present and future

Skew 2026, Padua (ITALY), January 7-9, 2026

Nicola Loperfido

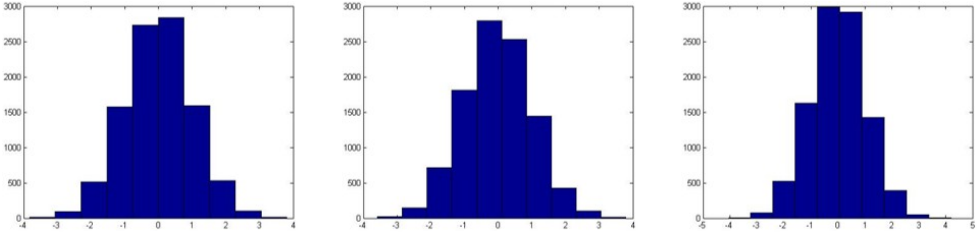
Università degli Studi di Urbino "Carlo Bo"
(ITALY)
nicola.loperfido@uniurb.it

January 9, 2026

Outline

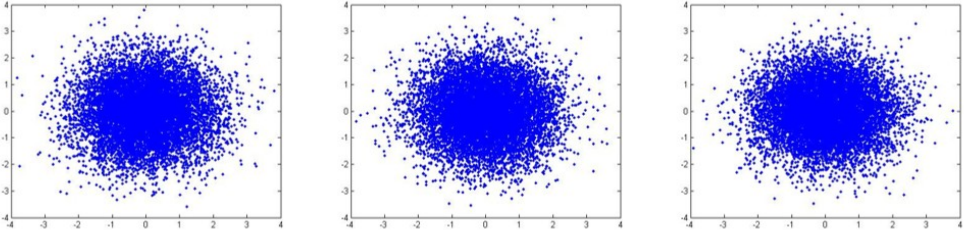
- Motivating example.
- Projection pursuit.
- Skewness optimization.
- Data example.
- Concluding remarks.

Motivating example: histograms



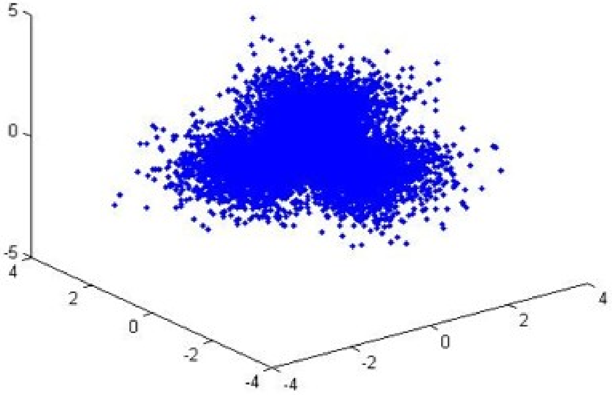
Each variable appears to be normal.

Motivating example: scatterplots



All scatterplots hint for bivariate standard normality.

Motivating example: 3-D plot



There seem to be three clusters.

Motivating example: distribution

The dataset is a random sample from the density

$$2\phi(z_1)\,\phi(z_2)\,\phi(z_3)\,\Phi(z_1 z_2 z_3)\,,\ \phi(\cdot) = \text{pdf of } N(0,1)\,,\ \Phi(\cdot) = \text{cdf of } N(0,1)\,.$$

$$\Downarrow$$

$$\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \sim \begin{pmatrix} Z_1 \\ Z_3 \end{pmatrix} \sim \begin{pmatrix} Z_2 \\ Z_3 \end{pmatrix} \sim N_2\left\{\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right\}.$$

Motivating example: remarks

- Data come from a trivariate, generalized skew-normal with standard normal bivariate marginals.
- The sampled distribution is standard: its mean and covariance are a null vector and an identity matrix.
- Neither subsets of the original variables nor principal components detect the data nonnormality.

Finding a data projection which retains interesting data features.

Projection pursuit: concept

Projection pursuit is a multivariate statistical technique aimed at finding interesting low-dimensional data projections.

Projection pursuit: definition

$I\left(X\right)=$ projection pursuit index evaluated at the random variable X.

$\mathbf{x}=$ p-dimensional random vector.

$\mathbb{R}_0^p=$ set of all nonnull p-dimensional real vectors.

$$\text{projecting direction} = \underset{\mathbf{c}\in\mathbb{R}_0^p}{\arg\max}\, I\left(\mathbf{c}^\top\mathbf{x}\right)$$

Projection pursuit: indices

- **Moments:** Skewness, kurtosis, ...

- **Distances:** Kolmogorov, Euclidean, ...

- **Information:** Shannon, Fisher, ...

Projection pursuit: default

- **Normality.** The normal distribution is the least interesting.

- **Invariance.** The projection index is affinely invariant.

- **Orthogonality.** The projecting directions are mutually orthogonal.

Projection pursuit: paradigm

- **Preprocessing** Remove data features which are either uninteresting, misleading or redundant.

- **Optimization** Optimize the projection pursuit index over all linear data projections, either globally or locally.

- **Removal** Remove from the data the detected interesting projections to detect other interesting projections.

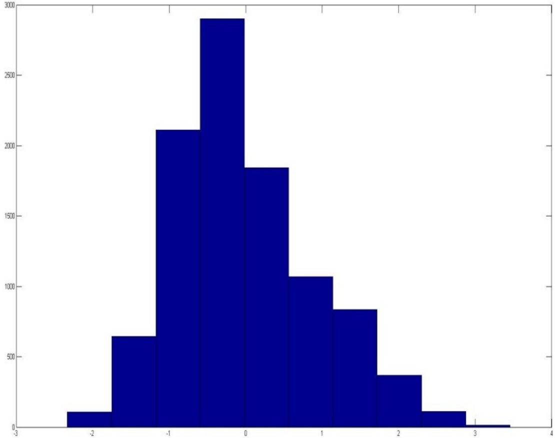
Projection pursuit: visualization

Projection pursuit supports the natural aptitude of the human vision at finding patterns in data graphical displays, thus bridging the gap between human and artificial intelligence.

Skewness optimization: definition

$$\beta_1^D\left(\mathbf{x}\right)=\max_{\mathbf{c}\in\mathbb{R}_0^p}\beta_1\left(\mathbf{c}^\top\mathbf{x}\right)=\max_{\mathbf{c}\in\mathbb{R}_0^p}\frac{\mathrm{E}^2\left\{\left(\mathbf{c}^\top\mathbf{x}-\mathbf{c}^\top\boldsymbol{\mu}\right)^3\right\}}{\left(\mathbf{c}^\top\boldsymbol{\Sigma}\mathbf{c}\right)^3}.$$

Skewness optimization: example



The projection which maximizes skewness clearly hints that data are skewed, but rightly does not suggest the presence of clusters.

Skewness optimization: models

- Skewed distributions.
- Finite mixtures.
- Independent components.

Skewness optimization: polynomial

Skewness optimization requires the optimization of a third order polynomial in several variables, subject to a quadratic constraint:

$$\beta_1^D(\mathbf{x}) = \max_{\mathbf{c}\in\mathbb{R}_0^p} \frac{\left\{\mathbf{c}^\top \boldsymbol{\Upsilon}(\mathbf{c}\otimes\mathbf{c})\right\}^2}{\left(\mathbf{c}^\top \boldsymbol{\Sigma}\mathbf{c}\right)^3},$$

$$\boldsymbol{\Upsilon} = cos(\mathbf{x}) = \mathrm{E}\left\{(\mathbf{x}-\boldsymbol{\mu})\otimes(\mathbf{x}-\boldsymbol{\mu})^\top\otimes(\mathbf{x}-\boldsymbol{\mu})^\top\right\}.$$

Skewness optimization: eigenpair

The directional skewness of a standardized random vector with finite third-order moments coincides with the squared dominant real tensor eigenvalue of its standardized coskewness, while the projecting direction is proportional to the associated tensor eigenvector:

$$\boldsymbol{\mu} = \mathrm{E}\left(\mathbf{z}\right) = \mathbf{0}_p,\ \boldsymbol{\Sigma} = \mathrm{E}\left(\mathbf{z}\mathbf{z}^\top\right) = \mathbf{I}_p,\ \boldsymbol{\Upsilon} = \mathrm{E}\left(\mathbf{z} \otimes \mathbf{z}^\top \otimes \mathbf{z}^\top\right),$$

$$\beta_1^D\left(\mathbf{z}\right) = \beta_1\left(\mathbf{c}_1^\top \mathbf{z}\right) = \lambda_1^2,\ \lambda_1 = \max_{\lambda \in \mathbb{R}} \left\{ \boldsymbol{\Upsilon}\left(\mathbf{c} \otimes \mathbf{c}\right) = \lambda \mathbf{c},\ \mathbf{c}^\top \mathbf{c} = 1 \right\}.$$

Skewness optimization: preprocessing

Let $\mathbf{x}$ and $\mathbf{v}$ be two independent and identically distributed random vectors with positive definite covariance matrices and finite third-order moments. The Mardia's skewness of $\mathbf{x}$ ($\mathbf{v}$) is

$$\beta_{1,p}(\mathbf{x}) = \mathrm{E}\left[\left\{(\mathbf{x}-\boldsymbol{\mu})\,\boldsymbol{\Sigma}^{-1}(\mathbf{v}-\boldsymbol{\mu})^{\top}\right\}^{3}\right].$$

Find a $k \times p$ full-rank matrix $\mathbf{A}$ minimizing the Mardia's skewness of $\mathbf{Ax}$, for a given value of $k < p$.

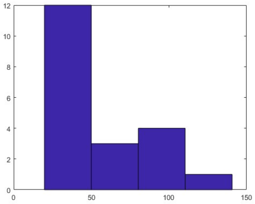
Data example: description

- Sources: Transfermarkt and Capology.

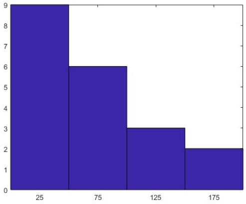
- Units: 20 major Italian football teams.

- Variables: Salaries, Acquisitions, Expenses.

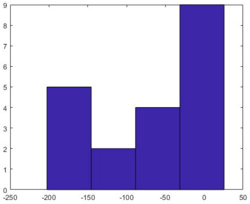
Data example: histograms



Salaries

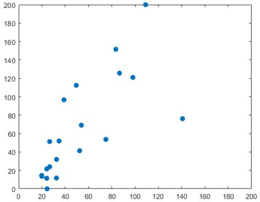


Acquisitions

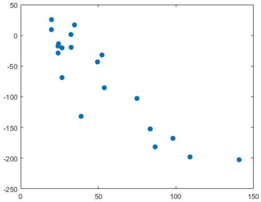


Expenses

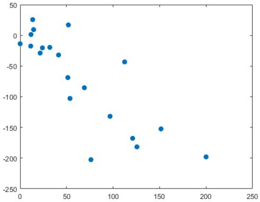
Data example: scatterplots



Salaries and Acquisitions

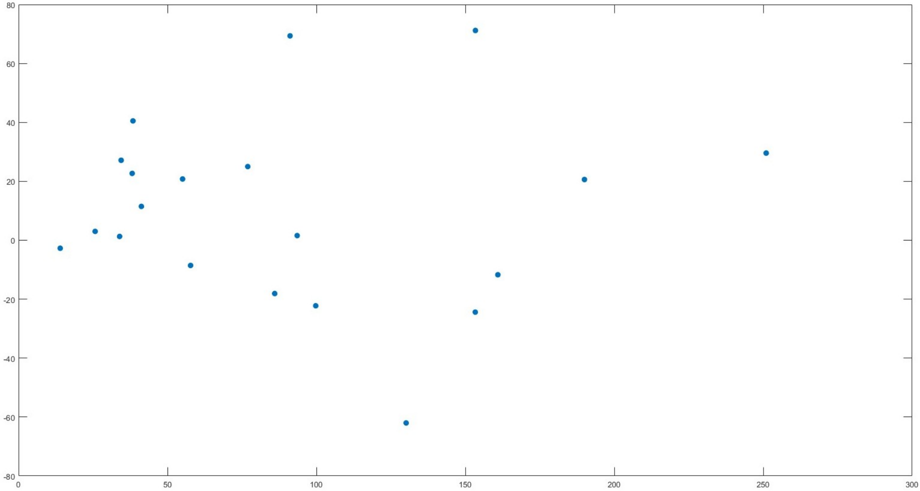


Salaries and Expenses



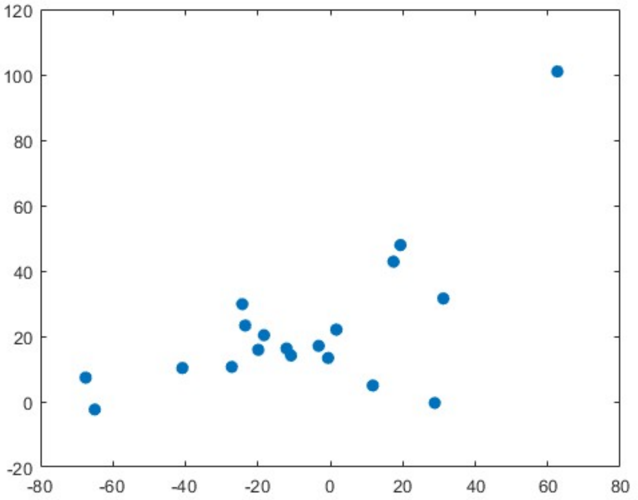
Acquisitions and Expenses

Data example: principal components



Scatterplot of the first and second principal components, accounting for more than 93% of the total variance.

Data example: skewness optimization



Scatterplot of the projections with minimal and maximal absolute skewnesses.

Data example: concluding remarks

- The scatterplot obtained via skewness optimization hints for the presence of an outlier: the Inter football team, which came second in the last Italian Football Championship.

- Other scatterplots obtained via projection pursuit, including the one based on kurtosis optimization (projections with minimal and maximal kurtoses), led to similar results.

- The principal component scatterplot did not suggest the presence of the potential outlier, despite accounting for about 93% of the total variance.

Future research: models

Skewness projection pursuit originated as an exploratory, data-based technique later has been successfully applied to (semi)parametric models.

- In which skewed models does SPP have a simple, interpretable form?
- Does SPP overcome some limitations of likelihood inference?
- Is there any scope for projection pursuit in Bayesian inference?

Future research: inference

The literature on the sampling properties of skewness projection pursuit (SPP) is quite limited.

- The sampling distribution of SPP is related to the maxima of Gaussian random fields.
- Sampling properties of SPP have been investigated using the tube method.
- Is the asymptotic distribution of directional skewness under normality skew-normal?

Future research: computation

The widespread use of (skewness) projection pursuit has been hampered by computational difficulties.

- Algorithms.
- Packages.
- Simulations.

Future research: eigenpairs

The literature on tensor eigenpairs is very limited and contains some wrong statements.

- Under which conditions tensor eigenpairs are real? .
- How many eigenpairs does a given tensor have?
- Assessing the multiplicity of tensor eigenvalues.

Future research: dimensionality

Projection pursuit poses additional problems when there are more variables than units.

- Data preprocessing for dimension reduction .
- Variable selection via projection pursuit.
- Generalized tensor eigenpairs to address singularities.

Future research: cumulants

The performance of SPP crucially depends on the structure of the multivariate third cumulant of the underlying distribution.

- Tensor rank of the third cumulant.
- Infinite third-order moments.
- Skewed distributions with null third cumulants.

References

- Franceschini, C. and Loperfido, N. (2017). MaxSkew, Orthogonal Data Projections with Maximal Skewness. R package version 1.1, https://CRAN.R- project.org/package=MaxSkew.
- Franceschini, C. and Loperfido, N. (2017). MultiSkew, Measures, Tests and Removes Multivariate Skewness. R package version 1.1.1, https://CRAN.R-project.org/package=MultiSkew.
- Friedman, J. H., and Tukey, J. W. (1974). A Projection Pursuit Algorithm for Exploratory Data Analysis. IEEE Transactions on Computers, Series C, 23, 881-890.
- Loperfido, N. (2024). Tensor Eigenvectors for Projection Pursuit. TEST 33, 453-472.
- Loperfido, N. and Shushi, T. (2025). Solving extended mean-variance models using tensor analysis. *The European Journal of Finance*, 1-11.
- Malkovich, J.F. and Afifi, A.A. (1973). On Tests for Multivariate Normality. *J. Amer. Statist. Ass.* **68**, 176-179.

THANK YOU FOR YOUR ATTENTION!