Why don't you put the whole paper in the title?

Andrew Byres, Camelot Diane, Eric Frederick and Sungkyu Jung
Seoul National University

November 26, 2020

Abstract

There should be a single paragraph summary which should not contain formulae or symbols, followed by some key words in alphabetical order. Typically there are 3–8 key words, which should contain nouns and be singular rather than plural. The summary contains bibliographic references only if they are essential. It should indicate results rather than describe the contents of the paper: for example, 'A simulation study is performed' should be replaced by a more informative phrase such as 'In a simulation our estimator had smaller mean square error than its main competitors.'

1 Introduction

This is a minimal instruction on writing a document in Latex. The end of a paragraph is marked in the .tex file by a blank line. Extra characters such as \\ at the end of lines or paragraphs should not be used.

References are added as needed. See the style in the following paragraph.

Candes and Tao (2007) used the Dantzig selector in high-dimensional estimation problems; see also Bickel et al. (2006); Bickel and Levina (2004). Chen and his colleagues pioneered the low-rank estimation principles in high dimensions (Chen and Buja, 2009; Chen and Qin, 2010; Chen and Cook, 2010).

2 Theorems and mathematical expressions

English sentences containing mathematical expressions or displayed formulae should be punctuated in the usual way: in particular please check carefully that all displayed expressions are correctly punctuated. Displayed expressions should be preceded by a colon only if grammatically warranted. Do not place a colon in the middle of a clause. See the following example.

The kurtosis of a random variable $X \sim f$ is defined as the fourth standard moment $\kappa(f) = \mathrm{E}(X-\mu)^4/\sigma^4$, where $\mu = \mathrm{E}(X)$, $\sigma^2 = \mathrm{Var}(X)$. Generalizing the univariate notion of kurtosis to multivariate distributions, Mardia (1970) defined the multivariate kurtosis $\kappa(f)$, for $\mathbf{X} \sim f$,

$$\kappa(f) = \mathrm{E}\{(\mathbf{X}-\mu)^\top \Sigma^{-1}(\mathbf{X}-\mu)\}^2, \tag{1}$$

where $\mu = \mathrm{E}(\mathbf{X})$ and $\Sigma = \mathrm{Var}(\mathbf{X})$. It is straightforward to check that the multivariate kurtosis is invariant under nonsingular affine transformations of the random vector $\mathbf{X}$. Thus, for the standardized random vector $\mathbf{Z} = \Sigma^{-1/2}(\mathbf{X}-\mu) \sim f_Z$, we have $\kappa(f) = \kappa(f_Z)$ for any f, and $\kappa(f_Z) = \mathrm{E}\{\mathbf{Z}^T\mathbf{Z}\}^2 = \mathrm{E}\{\|\mathbf{Z}\|^4\}$. In words, the multivariate kurtosis is the fourth moment of the distance of the standardized data points from the mean.

The multivariate kurtosis is easily computed from g:

Lemma 1. *For any rotationally symmetric density f on $\mathbb{R}^d$, let $g(r) = f((r,0,\ldots,0))$ and $c_g(p) = \int_0^\infty r^p g(r)dr$ for $p \geq 0$. Then,*

$$\kappa(f) = d^2 \frac{c_g(d+3)c_g(d-1)}{[c_g(d+1)]^2}. \tag{2}$$

It is a good idea to add an indicator of the theorem in the head of a proof.

Proof of Lemma 1. Suppose $\mathbf{X} \sim f$. Then, $\mathrm{Var}(\mathbf{X}) = \Sigma = \sigma^2 I$ for $\sigma^2 = d^{-1}\mathrm{trace}(\mathrm{Var}(\mathbf{X})) = d^{-1}\mathrm{E}\,\|X\|^2$. Then

$$\kappa(f) = \mathrm{E}(\mathbf{X}^T\Sigma^{-1}\mathbf{X})^2 = d^2\mathrm{E}\,\|\mathbf{X}\|^4/(\mathrm{E}\,\|\mathbf{X}\|^2)^2.$$

Note that for any $m \geq 0$, $\mathrm{E}\,\|\mathbf{X}\|^m = \mathrm{E}R^m$, for $R \sim f_R$. Pluggin in

$$\mathrm{E}R^m = \int_0^\infty r^m f_R(r)dr = \int_0^\infty \frac{1}{c_g(d-1)} r^{m+d-1} g(r)dr = \frac{c_g(d-1+m)}{c_g(d-1)}$$

and rearranging give the result. □

Use the align environment for multiple lines of equations. Equation numbers are given only if they are referenced.

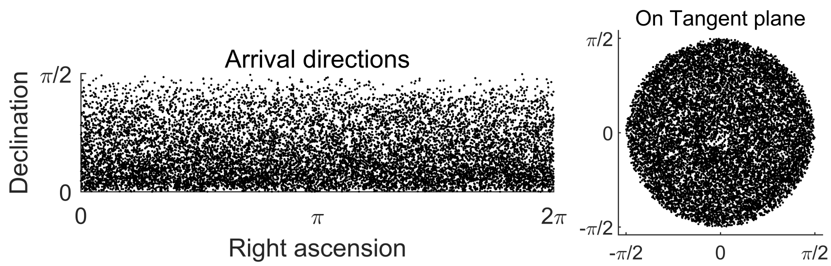

Figure 1: (left) The scatter of arrival directions of all $n = 12877$ cosmic events, shown in terms of right ascension (longitude) and declination (latitude). (right) The scatter of $\mathbf{y}_i = \mathrm{Log}_\mu \mathbf{x}_i$, where μ is set to the north pole.

$$
\begin{aligned}
\left\|\mathrm{Log}_\mu \mathbf{x} - \mathrm{Log}_{\hat{\mu}} \mathbf{x}\right\| &= \|\rho(\mu, \mathbf{x})\mathbf{v}(\mu) - \rho(\hat{\mu}, \mathbf{x})\mathbf{v}(\hat{\mu})\| \\
&= \|\rho(\mu, \mathbf{x})\mathbf{v}(\mu) - \rho(\hat{\mu}, \mathbf{x})\mathbf{v}(\mu) + \rho(\hat{\mu}, \mathbf{x})\mathbf{v}(\mu) - \rho(\hat{\mu}, \mathbf{x})\mathbf{v}(\hat{\mu})\| \\
&\leq |\rho(\mu, \mathbf{x}) - \rho(\hat{\mu}, \mathbf{x})| \, \|\mathbf{v}(\mu)\| + |\rho(\hat{\mu}, \mathbf{x})| \, \|\mathbf{v}(\mu) - \mathbf{v}(\hat{\mu})\| \\
&< \rho(\mu, \hat{\mu}) + 2\pi \sin(\epsilon/2) \\
&\leq c_0 \epsilon, \qquad (3)
\end{aligned}
$$

Do not begin a sentence with a symbol or number. For example, “Equation (3) shows that ...” not “(3) shows that ...”.

Remark 1. It can be checked that $\kappa(f)$ is invariant to the arbitrary scaling of the sectional density g.

Example 1. For reference, we compute $\kappa(f)$ for a few examples. The multivariate kurtosis of d-variate standard normal distribution is $d(d+2)$.

Examples and remarks can be references. See Example 1 and Remark 1.

3 Figures and tables

Figures should be appropriately placed, preferably at the top of pages. Figure 1 is an example data set. By an inspection of the left panel of Fig. 1, we conclude that the analysis was useless.

See also Tab. 1. Table 1 shows a typical style of formatting a data table.

Model	d	p	Allen	Thompson	Proposal	Heart	Love	Momocs
I	3	200	0.188	**0.150**	0.195	**0.154**	0.188	0.186
	3	500	0.208	**0.146**	0.219	0.152	0.208	0.208
	5	200	0.327	**0.221**	0.382	0.252	0.314	0.322
	5	500	0.352	**0.190**	0.409	0.324	0.337	0.370
II	3	200	**0.105**	0.165	0.156	0.161	0.123	**0.106**
	3	500	**0.107**	0.168	0.162	0.170	0.124	**0.108**
	5	200	**0.172**	0.489	0.326	0.463	0.242	0.243
	5	500	**0.182**	0.553	0.404	0.677	0.295	0.288
III	3	200	**0.116**	0.157	0.184	0.160	0.157	**0.119**
	3	500	**0.111**	0.156	0.190	0.161	0.161	**0.112**
	5	200	**0.214**	0.350	0.500	0.432	0.406	0.380
	5	500	**0.215**	0.343	0.606	0.552	0.490	0.461

Table 1: The minimal projection distance to the truth, averaged from 100 repetitions. The standard errors are at most 0.0206. Smaller distance indicates more precise estimation. Highlighted are the best performed models (within 2 standard error of the smallest).

4 한글

kotex 패키지를 이용하여 한글을 쓸 수 있다.

References

Bickel, P. and Levina, E. (2004), "Some theory for Fisher's linear discriminant function , 'naive Bayes', and some alternatives when there are many more variables than observations," *Bernoulli*, 10, 989–1010.

Bickel, P., Li, B., Tsybakov, A., van de Geer, S., Yu, B., Valdés, T., Rivero, C., Fan, J., and van der Vaart, A. (2006), "Regularization in statistics," *Test*, 15, 271–344.

Candes, E. and Tao, T. (2007), "The Dantzig selector : statistical estimation when p is much larger than n," *Annals of Statistics*, 35, 2313–2351.

Chen, L. and Buja, A. (2009), "Local Multidimensional Scaling for Nonlinear Dimension Reduction, Graph Drawing, and Proximity Analysis," *Journal of the American Statistical Association*, 104, 209–219.

Chen, S. X. and Qin, Y.-L. (2010), "A two-sample test for high-dimensional data with applications to gene-set testing," *The Annals of Statistics*, 38, 808–835.

Chen, X. and Cook, R. D. (2010), "Some insights into continuum regression and its asymptotic properties," *Biometrika*, 97, 985–989.

Mardia, K. V. (1970), "Measures of multivariate skewness and kurtosis with applications," *Biometrika*, 57, 519–530.