Journal of Statistical Planning and Inference 143 (2013) 2016–2022

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

A note on Stein's lemma for multivariate elliptical distributions

Zinoviy Landsman [a], Steven Vanduffel [b], Jing Yao [c,*,1]

[a] Department of Statistics, University of Haifa, Haifa, 31905, Israel
[b] Faculty of Economics, Vrije Universiteit Brussel (VUB), Pleinlaan 2, BE-1050 Brussels, Belgium
[c] School of Insurance, Southwestern University of Finance and Economics, Chengdu, 610074, China

ARTICLE INFO

Article history:
Received 20 September 2012
Received in revised form
27 May 2013
Accepted 1 June 2013
Available online 22 June 2013

Keywords:
Conditional expectation
Stein's lemma
Multivariate elliptical distribution
Siegel's formula

ABSTRACT

When two random variables are bivariate normally distributed Stein's original lemma allows to conveniently express the covariance of the first variable with a function of the second. Landsman and Neslehova (2008) extend this seminal result to the family of multivariate elliptical distributions. In this paper we use the technique of conditioning to provide a more elegant proof for their result. In doing so, we also present a new proof for the classical linear regression result that holds for the elliptical family.

1. Introduction

Stein's (1973) lemma demonstrates that for a bivariate normal random vector (X,Y), it holds that

$$\mathrm{Cov}[X,h(Y)]=\mathrm{Cov}[X,Y]\mathrm{E}[h'(Y)], \tag{1}$$

where $h(y)$ is any differentiable function such that $\mathrm{E}[|h'(Y)|]$ exists. Since the original work of Stein, several generalizations appeared in the literature. Stein (1981) and then Liu (1994) generalized the lemma to a multivariate normal context. Further, Landsman (2006) showed that Stein's type lemma also holds when (X,Y) is bivariate elliptically distributed, and Landsman and Neslehova (2008) extended this result to multivariate elliptical vectors. Elliptical random vectors do not exhibit skewness and this feature motivated Adcock (2007) to nicely derive a version of Stein's lemma when the vector of random variables has a multivariate skew-normal distribution; see also Adcock (2010). The technical efforts carried out to generalize Stein's original work are not a surprise. Indeed, the lemma has shown to be very useful in various fields including statistics, probability, decision theory, finance and insurance. Most of these applications were initially derived in a multivariate normal context but the generalizations that appeared in the literature usually allow to extend these results to broader and, in many cases, more realistic settings. We now describe some of the applications in some more length: *First*, when the returns are multivariate normally distributed the implication of Stein's lemma is that all rational investors will select a portfolio which lies on Markowitz' mean-variance efficient frontier. In other words, Stein's lemma greatly simplifies the task of portfolio selection. *Second*, Liu (1994) shows how Stein's lemma allows to prove and to extend Siegel's (1993) formula for the covariance of an arbitrary element of a multivariate normal vector with its minimum element. This last result is useful to optimize hedging schemes for commodity contracts. *Third*, the lemma has also gained interest in the field of financial

* Corresponding author.
E-mail addresses: landsman@stat.haifa.ac.il (Z. Landsman), steven.vanduffel@vub.ac.be (S. Vanduffel), jingyao@vub.ac.be (J. Yao).
[1] This research is carried out when the author is affiliated to Vrije Universiteit Brussel (VUB).

0378-3758/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jspi.2013.06.003

economics and in insurance for its important application to Capital Asset Pricing Models (CAPMs). In particular, the classical CAPM result expressed as $\mathrm{E}[R_k]=R_F+\beta(\mathrm{E}[R_M]-R_F)$ gives the expected return $\mathrm{E}[R_k]$ of the k-th asset as a linear function of the fixed risk-free rate R_F and the expected return $\mathrm{E}[R_M]$ of the market portfolio; see Panjer (1998, Section 4.5) and alsoCochrane (2001). *Fourth*, in statistics Stein's lemma appears to be useful for improved estimation of $\mathrm{Cov}[X,h(Y)]$ (Landsman and Neslehova (2008)), as well as exact moment formulas and Bayesian statistics (Brown et al., 2006). These authors use the heat equation to establish an identity which is closely related to Stein's lemma. They also provide and discuss a series of other applications ranging from an improvement of Jensen's celebrated inequality, the establishment of inadmissibility results in decision theory and the counting of matchings in graphs in the field of graph theory.

In this paper we make the following contributions. Landsman and Neslehova (2008) have provided us with a version of Stein's lemma, valid for the family of general multivariate elliptical distributions. Our *first contribution* is that we provide a simpler and more instructive proof for their result. The key difference between their approach and ours, allowing us to simplify the proof considerably, is that by conditioning we can take advantage of the classic linear regression result that holds for the elliptical family. This result essentially states that when (X,Y) is bivariate elliptically distributed, the best (with respect to the L^2-norm $\|\cdot\|$) predictor for Y based on X is the appropriate linear function of X. Our *second contribution* is that we give a new pedagogical proof for this seminal regression result. Finally, we also show that Liu's (1994) generalization of Siegel's (1993) covariance formula also holds true in a multivariate elliptical context.

The paper is organized as follows. Section 2 provides definitions and preliminary results. Section 3 introduces the multivariate elliptical distributions and provides our proof for the linear regression result. Next in Section 4 we present our alternative proof for Stein's lemma in a multivariate elliptical context. Section 5 provides the generalization of Siegel's (1993) covariance formula and final remarks are presented in Section 6.

2. Preliminaries

Let $L^2(\mathbb{P})$ be the space of square integrable random variables, which are all defined on a given probability space $(\Omega,\mathcal{F},\mathbb{P})$. We equip $L^2(\mathbb{P})$ with the inner product $\langle X,Y\rangle=\mathrm{E}[XY]$ with an associated norm denoted by $\|\cdot\|$. It is then well-known that $L^2(\mathbb{P})$ is a Hilbert space. Consider now linearly independent random variables $1,X_1,X_2,\dots,X_n\in L^2(\mathbb{P})$. They can be used to generate a *linear* subspace H. Then for any $Y\in L^2(\mathbb{P})$ we have that its projection $P(Y)$ on the subspace H is given by

$$P(Y)=\mathrm{E}[Y]+\lambda^T(\mathbf{X}-\mathrm{E}[\mathbf{X}]). \tag{2}$$

Here, $\mathbf{X}^T:=(X_1X_2,\dots,X_n)$ and $\lambda^T:=(\lambda_1,\lambda_2,\dots,\lambda_n)$ is the vector of coefficients given by

$$\lambda=(\mathrm{Cov}[\mathbf{X}])^{-1}\mathrm{Cov}[\mathbf{X},Y],$$

where the $n\times n$ matrix $\mathrm{Cov}[\mathbf{X}]$ (called the covariance matrix of $\mathbf{X}$) and the $n\times 1$ matrix $\mathrm{Cov}[\mathbf{X},\mathrm{Y}]$ are respectively defined through

$$(\mathrm{Cov}[\mathbf{X}])_{i,j}:=\mathrm{Cov}[X_i,X_j],\quad i,j=1,2,\dots,n,$$
$$(\mathrm{Cov}[\mathbf{X},Y])_i:=\mathrm{Cov}[X_i,Y],\quad i=1,2,\dots,n,$$

see also Kutner et al. (2004). Note that for a given $Y\in L^2(\mathbb{P})$ it holds for all $Z\in H$ that $\|P(Y)-Y\|\le\|Z-Y\|$. Hence, $P(Y)$ can also be interpreted as the best possible linear combination of the random variables $1, X_1,X_2,\dots,X_n$ in the sense that the L^2-distance with Y becomes minimal. In other words, $P(Y)$ is the best linear predictor of Y based on the random variables $1,X_1,X_2,\dots,X_n$.

Let us next aim at finding a more general Borel function f: $\mathbb{R}^n\to\mathbb{R}$ such that now $\|f(1,X_1,X_2,\dots,X_n)-Y\|$ becomes minimal. To this end, we consider the probability space $(\Omega,\sigma(\mathbf{X}),\mathbb{P})$ where $\sigma(\mathbf{X})$ is the smallest algebra such that the random variables $X_1,X_2,\dots,X_n$ are measurable (note that constant functions are always measurable). Let us also denote by F the corresponding space of square integrable random variables, and we equip it with the same inner product as $L^2(\mathbb{P})$. It is then clear that $H\subseteq F\subseteq L^2(\mathbb{P})$. It is also known that the projection of Y on F is given by $\mathrm{E}[Y|\sigma(\mathbf{X})]$. In the literature, with some abuse of notation, one often writes $\mathrm{E}[Y|\mathbf{X}]$ instead of $\mathrm{E}[Y|\sigma(\mathbf{X})]$.

3. Multivariate elliptical distributions

Definition 1 (*Multivariate elliptical distribution*). The random vector $\mathbf{X}=(X_1,\dots,X_n)^T$ has an elliptical distribution with parameters the $n\times 1$ vector $\boldsymbol{\mu}$ and the $n\times n$ positive definite matrix $\boldsymbol{\Sigma}$ if its characteristic function is given by

$$\mathrm{E}[\exp(i\mathbf{t}^T\mathbf{X})]=\exp(i\mathbf{t}^T\boldsymbol{\mu})\phi\left(\frac{\mathbf{t}^T\boldsymbol{\Sigma}t}{2}\right),\quad \mathbf{t}^T=(t_1,t_2,\dots,t_n), \tag{3}$$

for some scalar function $\phi(t)$ which is called the characteristic generator. We then write $\mathbf{X}\sim E_n(\boldsymbol{\mu},\boldsymbol{\Sigma},\phi)$.

A necessary and sufficient condition for the function ϕ to be a characteristic generator of an n-dimensional elliptical distribution is given in Theorem 2.2 of Fang et al. (1990). In general, the ellipsoidal vector $\mathbf{X}\sim E_n(\boldsymbol{\mu},\boldsymbol{\Sigma},\phi)$ may not have a