Elect. Comm. in Probab. **11** (2006), 149–159

ELECTRONIC COMMUNICATIONS in PROBABILITY

LARGE AND MODERATE DEVIATIONS FOR HOTELLING'S T^2-STATISTIC

AMIR DEMBO [1]
Department of Statistics, Stanford University, Stanford, CA 94305-4065, USA
email: amir@stat.stanford.edu

QI-MAN SHAO [2]
Department of Mathematics, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China; Department of Mathematics, University of Oregon, Eugene, OR 97403, USA; Department of Mathematics, Zhejiang University, Hangzhou, Zhejiang 310027, China.
email: maqmshao@ust.hk

Submitted February 7 2006, *accepted in final form* July 14 2006

AMS 2000 Subject classification: Primary 60F10, 60F15, secondly 62E20, 60G50
Keywords:
large deviation; moderate deviation; self-normalized partial sums; law of the iterated logarithm; T^2 statistic.

Abstract
Let $\boldsymbol{X}, \boldsymbol{X}_1, \boldsymbol{X}_2, ...$ be i.i.d. R^d-valued random variables. We prove large and moderate deviations for Hotelling's T^2-statistic when $\boldsymbol{X}$ is in the generalized domain of attraction of the normal law.

1 Introduction

Let $\boldsymbol{X}, \boldsymbol{X}_1, \boldsymbol{X}_2, ...$ be a sequence of independent and identically distributed (i.i.d.) nondegenerate R^d-valued random vectors with mean $\boldsymbol{\mu}$, where $d \geq 1$. Let

$$\boldsymbol{S}_n = \sum_{i=1}^n \boldsymbol{X}_i, \ \boldsymbol{V}_n = \sum_{i=1}^n (\boldsymbol{X}_i - \boldsymbol{S}_n/n)(\boldsymbol{X}_i - \boldsymbol{S}_n/n)'$$

Define Hotelling's T^2 statistic by

$$T_n^2 = (\boldsymbol{S}_n - n\boldsymbol{\mu})' \boldsymbol{V}_n^{-1} (\boldsymbol{S}_n - n\boldsymbol{\mu}). \tag{1.1}$$

[1] RESEARCH PARTIALLY SUPPORTED BY NSF GRANTS #DMS-0406042 AND #DMS-FRG-0244323
[2] RESEARCH PARTIALLY SUPPORTED BY DAG05/06. SC27 AT HKUST

The T^2-statistic is used for testing hypotheses about the mean $\boldsymbol{\mu}$ and for obtaining confidence regions for the unknown $\boldsymbol{\mu}$. When $\boldsymbol{X}$ has a normal distribution $N(\boldsymbol{\mu}, \Sigma)$, it is known that $(n-d)T_n^2/(dn)$ is distributed as an F-distribution with d and $n-d$ degrees of freedom (see, e.g., Anderson (1984)). The T^2-test has a number of optimal properties. It is uniformly most powerful in the class of tests whose power function depends only on $\boldsymbol{\mu}'\Sigma^{-1}\boldsymbol{\mu}$ (Simaika (1941)), is admissible (Stein (1956) and Kiefer and Schwartz (1965)), and is robust (Kariya (1981)). One can refer to Muirhead (1982) for other invariant properties of the T^2-test. When the distribution of $\boldsymbol{X}$ is not normal, it was proved by Sepanski (1994) that the limiting distribution of T_n^2 as $n \to \infty$ is a χ^2-distribution with d degrees of freedom. An asymptotic expansion for the distribution of T_n^2 is obtained by Fujikoshi (1997) and Kano (1995) independently. The main aim of this note is to give a large and moderate deviations for the T^2-statistic.

THEOREM **1.1** *Assume that $\boldsymbol{\mu} = 0$. For $\alpha \in (0,1)$, let*

$$K(\alpha) = \sup_{b\geq 0} \sup_{||\boldsymbol{\theta}||=1} \inf_{t\geq 0} E\exp\Big(t\big(b\boldsymbol{\theta}'\boldsymbol{X} - \alpha((\boldsymbol{\theta}'\boldsymbol{X})^2 + b^2)/2\big)\Big). \tag{1.2}$$

Then, for all $x > 0$,

$$\lim_{n\to\infty} P\Big(T_n^2 \geq x\, n\Big)^{1/n} = K(\sqrt{x/(1+x)}). \tag{1.3}$$

THEOREM **1.2** *Let $\{x_n,\ n \geq 1\}$ be a sequence of positive numbers with $x_n \to \infty$ and $x_n = o(n)$ as $n \to \infty$. Assume that $h(x) := E||\boldsymbol{X}||^2 1\{||\boldsymbol{X}|| \leq x\}$ is slowly varying and*

$$\liminf_{x\to\infty} \inf_{\boldsymbol{\theta}\in R^d, ||\boldsymbol{\theta}||=1} E(\boldsymbol{\theta}'\boldsymbol{X})^2 1\{||\boldsymbol{X}|| \leq x\}/h(x) > 0\ . \tag{1.4}$$

If $\boldsymbol{\mu} = 0$, then

$$\lim_{n\to\infty} x_n^{-1} \ln P\Big(T_n^2 \geq x_n\Big) = -\frac{1}{2}. \tag{1.5}$$

From Theorem **??** we have the following law of the iterated logarithm.

THEOREM **1.3** *Assume that $h(x) := E||\boldsymbol{X}||^2 1\{||\boldsymbol{X}|| \leq x\}$ is slowly varying and* (**??**) *is satisfied. If $\boldsymbol{\mu} = 0$, then*

$$\limsup_{n\to\infty} \frac{T_n^2}{2\log\log n} = 1 \quad a.s.$$

Theorems **??** and **??** demonstrate again that the Hotelling's T^2 statistic is very robust. Theorem **??** also provides a direct tool to estimate the efficiency of the T^2 test, such as the Bahadur efficiency. See He and Shao (1996).

Theorems **??** and **??** are in the context of the so-called self-normalized limit theorems. The past decade has witnessed important developments in this area. One can refer to Griffin and Kuelbs (1989) for the self-normalized law of the iterated logarithm when $d = 1$; Dembo and Shao (1998a, 1998b) for $d \geq 1$; Shao (1997) for self-normalized large and moderate deviations of i.i.d. sums; Faure (2002) for self-normalized large deviation for Markov chains; Jing, Shao and Zhou (2004) for self-normalized saddlepoint approximation; Jing, Shao and Wang (2003) for self-normalized Cramér- type large deviations for independent random variables; Bercu, Gassiat and Rio (2002) for large and moderate deviations for self-normalized empirical processes; Chistyakov and Götze (2004a) for the necessary and sufficient condition for having a non-degenerate limiting distribution of self-normalized sums; Shao (1998, 2004) for surveys of recent developments in this subject. Other self-normalized large deviation results can be found in Chistyakov and Götze (2004b), Robinson and Wang (2004) and Wang (2005).

REMARK **1.1** *Following Dembo and Shao (1998b), it is possible to have a large deviation principle for T_n^2. Formula (1.2) may become clearer from the large deviation principle point of view. However, it may be not easy to compute $K(\alpha)$ in general.*

REMARK **1.2** *It is easy to see that when $E||\boldsymbol{X}||^2 < \infty$ and $\boldsymbol{X}$ is nondegenerate, $h(x)$ converges to a constant and (**??**) is satisfied.*

REMARK **1.3** *In (**??**) when $\boldsymbol{V}_n$ is not full rank, i.e., $\boldsymbol{V}_n$ is degenerate, $\boldsymbol{x}'\boldsymbol{V}_n^{-1}\boldsymbol{x}$ is defined as (see (**??**) in the next section)*

$$\boldsymbol{x}'\boldsymbol{V}_n^{-1}\boldsymbol{x} = \sup_{||\boldsymbol{\theta}||=1, \boldsymbol{\theta}'\boldsymbol{x}\geq 0} \frac{(\boldsymbol{\theta}'\boldsymbol{x})^2}{\boldsymbol{\theta}'\boldsymbol{V}_n\boldsymbol{\theta}},$$

*where $0/0$ is interpreted as ∞. The latter convention is the reason why $b = 0$ is allowed in the definition (**??**) of $K(\alpha)$, which is essential for the validity of Theorem **??** in case the law of $\boldsymbol{X}$ has atoms.*

REMARK **1.4** *$\boldsymbol{X}$ is said to be in the generalized domain of attraction of the normal law ($\boldsymbol{X} \in$ GDOAN) if there exist nonrandom matrices $\boldsymbol{A}_n$ and constant vector $\boldsymbol{b}_n$ such that*

$$\boldsymbol{A}_n(\boldsymbol{S}_n - \boldsymbol{b}_n) \stackrel{d.}{\to} N(0, \boldsymbol{I}).$$

Hahn and Klass (1980) proved that $\boldsymbol{X} \in GDOAN$ if and only if

$$\lim_{x\to\infty} \sup_{||\boldsymbol{\theta}||=1} \frac{x^2 P(|\boldsymbol{\theta}'\boldsymbol{X}| > x)}{E|\boldsymbol{\theta}'\boldsymbol{X}|^2 I\{|\boldsymbol{\theta}'\boldsymbol{X}| \leq x\}} = 0. \tag{1.6}$$

*If conditions in Theorem **??** are satisfied, then (**??**) holds. We conjecture that Theorem **??** remains valid under condition (**??**).*

2 Proofs

Let $\boldsymbol{B}$ be an $d \times d$ symmetric positive definite matrix. Then, clearly,

$$\begin{aligned}\forall\ \boldsymbol{x} \in R^d,\ \ \boldsymbol{x}'\boldsymbol{B}^{-1}\boldsymbol{x} &= \sup_{\boldsymbol{\vartheta}\in R^d} \Big(2\boldsymbol{\vartheta}'\boldsymbol{x} - \boldsymbol{\vartheta}'\boldsymbol{B}\boldsymbol{\vartheta}\Big) = \sup_{||\boldsymbol{\theta}||=1, b\geq 0} \Big\{2b\boldsymbol{\theta}'\boldsymbol{x} - b^2\boldsymbol{\theta}'\boldsymbol{B}\boldsymbol{\theta}\Big\} \\ &= \sup_{||\boldsymbol{\theta}||=1, \boldsymbol{\theta}'\boldsymbol{x}\geq 0} \frac{(\boldsymbol{\theta}'\boldsymbol{x})^2}{\boldsymbol{\theta}'\boldsymbol{B}\boldsymbol{\theta}}\end{aligned} \tag{2.1}$$

(taking $\boldsymbol{\vartheta} = b\boldsymbol{\theta}$, with $b \geq 0$ and $||\boldsymbol{\theta}|| = 1$).
Proof of Theorem ??. Letting

$$\boldsymbol{\Gamma}_n = \sum_{i=1}^n \boldsymbol{X}_i\boldsymbol{X}_i',$$

we can rewrite $\boldsymbol{V}_n$ as

$$\boldsymbol{V}_n = \boldsymbol{\Gamma}_n - \boldsymbol{S}_n \boldsymbol{S}_n'/n.$$

By (**??**), for any $a > 0$

$$\begin{aligned} \{T_n^2 \geq a^2\} &= \{\boldsymbol{S}_n' \boldsymbol{V}_n^{-1} \boldsymbol{S}_n \geq a^2\} \\ &= \Big\{\exists\, \boldsymbol{\theta} \in R^d, ||\boldsymbol{\theta}|| = 1, \boldsymbol{\theta}' \boldsymbol{S}_n / \sqrt{\boldsymbol{\theta}' \boldsymbol{V}_n \boldsymbol{\theta}} \geq a\Big\} \\ &= \Big\{\exists\, \boldsymbol{\theta} \in R^d, ||\boldsymbol{\theta}|| = 1, \boldsymbol{\theta}' \boldsymbol{S}_n \geq a \sqrt{\boldsymbol{\theta}' \boldsymbol{\Gamma}_n \boldsymbol{\theta} - (\boldsymbol{\theta}' \boldsymbol{S}_n)^2/n}\Big\} \\ &= \Big\{\exists\, \boldsymbol{\theta} \in R^d, ||\boldsymbol{\theta}|| = 1, \boldsymbol{\theta}' \boldsymbol{S}_n \geq \frac{a}{\sqrt{1 + a^2/n}} \sqrt{\boldsymbol{\theta}' \boldsymbol{\Gamma}_n \boldsymbol{\theta}}\Big\} \end{aligned} \tag{2.2}$$

Hence, for all $x > 0$

$$P\Big(T_n^2 \geq x\, n\Big) = P\Big(\sup_{||\boldsymbol{\theta}||=1} \frac{\boldsymbol{\theta}' \boldsymbol{S}_n}{\sqrt{\boldsymbol{\theta}' \boldsymbol{\Gamma}_n \boldsymbol{\theta}}} \geq (x/(1+x))^{1/2} n^{1/2}\Big) \tag{2.3}$$

Notice that

$$\boldsymbol{\theta}' \boldsymbol{S}_n = \sum_{i=1}^n \boldsymbol{\theta}' \boldsymbol{X}_i \quad \text{and} \quad \boldsymbol{\theta}' \boldsymbol{\Gamma}_n \boldsymbol{\theta} = \sum_{i=1}^n (\boldsymbol{\theta}' \boldsymbol{X}_i)^2$$

By Theorem 1.1 of Shao (1997), it follows from (**??**) that

$$\liminf_{n\to\infty} P\Big(T_n^2 \geq x\, n\Big)^{1/n} \geq K(\sqrt{x/(x+1)})$$

(for $K(\cdot)$ of (**??**)). To prove the upper bound of (**??**), it suffices to show that for $\alpha \in (0,1)$

$$\limsup_{n\to\infty} P\Big(\sup_{||\boldsymbol{\theta}||=1} \Big\{\boldsymbol{\theta}' \boldsymbol{S}_n - \alpha n^{1/2} \sqrt{\boldsymbol{\theta}' \boldsymbol{\Gamma}_n \boldsymbol{\theta}}\Big\} \geq 0\Big)^{1/n} \leq K(\alpha). \tag{2.4}$$

Let $A \geq 2$ and define $\xi_i(\boldsymbol{\theta}) := \xi_i(\boldsymbol{\theta}, A) = \boldsymbol{\theta}' \boldsymbol{X}_i 1\{||\boldsymbol{X}_i|| \leq A\}$. We can make the proof of the upper bound with any fixed $\alpha \in (0,1)$ and $\varepsilon \in (0, 1/2)$,

$$\begin{aligned} & P\Big(\sup_{||\boldsymbol{\theta}||=1} \Big\{\boldsymbol{\theta}' \boldsymbol{S}_n - \alpha n^{1/2} \sqrt{\boldsymbol{\theta}' \boldsymbol{\Gamma}_n \boldsymbol{\theta}}\Big\} \geq 0\Big) \\ \leq\ & P\Big(\sup_{||\boldsymbol{\theta}||=1} \Big\{\sum_{i=1}^n \xi_i(\boldsymbol{\theta}) - (1-\varepsilon)\alpha n^{1/2} (\sum_{i=1}^n \xi_i^2(\boldsymbol{\theta}))^{1/2}\Big\} \geq 0\Big) \\ +\ & P\Big(\sup_{||\boldsymbol{\theta}||=1} \Big\{\sum_{i=1}^n \boldsymbol{\theta}' \boldsymbol{X}_i 1\{||\boldsymbol{X}_i|| > A\} - \varepsilon \alpha n^{1/2} (\sum_{i=1}^n (\boldsymbol{\theta}' \boldsymbol{X}_i)^2)^{1/2}\Big\} \geq 0\Big) \\ :=\ & I_1 + I_2. \end{aligned} \tag{2.5}$$

By the Cauchy inequality and

$$\forall\ \ a > 0, \quad P(B(n,p) \geq an) \leq (3p/a)^{an} \tag{2.6}$$

for the binomial random variable $B(n,p)$, we have

$$\begin{aligned} \limsup_{n\to\infty} I_2^{1/n} &\leq \limsup_{n\to\infty} P\Big(\sum_{i=1}^n 1\{||\boldsymbol{X}_i|| > A\} \geq (\varepsilon\alpha)^2 n\Big) \\ &\leq (3(\alpha\varepsilon)^{-2} P(||\boldsymbol{X}|| > A))^{(\alpha\varepsilon)^2}. \end{aligned} \tag{2.7}$$

It remains to bound I_1. Using the representation

$$\forall\ y>0,\ x\geq 0,\ z\geq x/y \qquad x\,y=(1/2)\inf_{0<b\leq z}\frac{1}{b}\Big(x^2+b^2y^2\Big),$$

we see that

$$(\sum_{i=1}^n \xi_i^2(\boldsymbol{\theta}))^{1/2}n^{1/2}=(1/2)\inf_{0<b\leq A}\frac{1}{b}\Big(\sum_{i=1}^n\xi_i^2(\boldsymbol{\theta})+b^2n\Big)$$

and

$$\begin{aligned} I_1 &= P\Big(\bigcup_{||\boldsymbol{\theta}||=1}\Big\{\sum_{i=1}^n\xi_i(\boldsymbol{\theta})\geq\frac{(1-\varepsilon)\alpha}{2}\inf_{0<b\leq A}\frac{1}{b}\Big(\sum_{i=1}^n\xi_i^2(\boldsymbol{\theta})+b^2n\Big)\Big\}\Big) \qquad (2.8)\\ &= P\Big(\sup_{0\leq b\leq A}\sup_{||\boldsymbol{\theta}||=1}\sum_{i=1}^n Z_i(\boldsymbol{\theta},b)\geq 0\Big), \end{aligned}$$

where $Z_i(\boldsymbol{\theta},b):=b\xi_i(\boldsymbol{\theta})-(1-\varepsilon)\alpha(\xi_i^2(\boldsymbol{\theta})+b^2)/2$. Let $0<\eta<1/4$ and consider a finite η-cover $\mathcal{G}$ of $\{(\boldsymbol{\theta},b):\ \boldsymbol{\theta}\in R^d, ||\boldsymbol{\theta}||=1, 0\leq b\leq A\}$ with respect to maximum norm in R^{d+1}. That is, for any $0\leq b\leq A$ and $\boldsymbol{\theta}\in R^d$ with $||\boldsymbol{\theta}||=1$, there exists $(\boldsymbol{\theta}_0,b_0)\in\mathcal{G}$ such that

$$||\boldsymbol{\theta}-\boldsymbol{\theta}_0||_\infty\leq\eta,\ |b-b_0|\leq\eta,\ \text{and}\ \ ||\boldsymbol{\theta}_0||=1\ . \qquad (2.9)$$

Since $|\xi_i(\boldsymbol{\theta})|\leq A$ it follows that for some $C=C(\alpha,d)<\infty$ all i and all $(\boldsymbol{\theta}_0,b_0)\in\mathcal{G}$,

$$\sup_{|b-b_0|\leq\eta}\ \sup_{||\boldsymbol{\theta}-\boldsymbol{\theta}_0||_\infty\leq\eta}|Z_i(\boldsymbol{\theta},b)-Z_i(\boldsymbol{\theta}_0,b_0)|\leq CA^2\eta\,. \qquad (2.10)$$

By Chebyshev's inequality we obtain that

$$\begin{aligned} & P\Big(\sup_{|b-b_0|\leq\eta}\ \sup_{||\boldsymbol{\theta}-\boldsymbol{\theta}_0||_\infty\leq\eta}\sum_{i=1}^n Z_i(\boldsymbol{\theta},b)\geq 0\Big) \qquad (2.11)\\ &\leq \inf_{t\geq 0}\Big\{e^{tCA^2\eta}E\exp(tZ(\boldsymbol{\theta}_0,b_0))\Big\}^n\\ &\leq \inf_{0\leq t\leq m}\Big\{e^{tCA^2\eta}E\exp(tZ(\boldsymbol{\theta}_0,b_0))\Big\}^n \end{aligned}$$

for any $m>0$, where $Z(\boldsymbol{\theta},b):=b\boldsymbol{\theta}'\boldsymbol{X}1\{||\boldsymbol{X}||\leq A\}-(1-\varepsilon)\alpha((\boldsymbol{\theta}'\boldsymbol{X})^21\{||\boldsymbol{X}||\leq A\}+b^2)/2$. Hence,

$$\limsup_{n\to\infty} I_1^{1/n}\leq\sup_{0\leq b\leq A}\ \sup_{||\boldsymbol{\theta}||=1}\ \inf_{0\leq t\leq m}\ e^{tCA^2\eta}E\exp(tZ(\boldsymbol{\theta},b))\ . \qquad (2.12)$$

Let $V(\boldsymbol{\theta},b,\varepsilon):=b\boldsymbol{\theta}'\boldsymbol{X}-(1-\varepsilon)\alpha((\boldsymbol{\theta}'\boldsymbol{X})^2+b^2)/2$. Then, for all $t\geq 0$,

$$E\exp(tZ(\boldsymbol{\theta},b))\leq E\exp(tV(\boldsymbol{\theta},b,\varepsilon))+P(||\boldsymbol{X}||>A)\,.$$

Therefore, considering $\eta\downarrow 0$ and then $A\uparrow\infty$, it follows from (**??**), (**??**) and (**??**) that

$$\begin{aligned} &\limsup_{n\to\infty}P\Big(\sup_{||\boldsymbol{\theta}||=1}\frac{\boldsymbol{\theta}'\boldsymbol{S}_n}{\sqrt{\boldsymbol{\theta}'\boldsymbol{\Gamma}_n\boldsymbol{\theta}}}\geq\alpha n^{1/2}\Big)^{1/n}\\ &\leq \sup_{b\geq 0}\ \sup_{||\boldsymbol{\theta}||=1}\ \inf_{0\leq t\leq m}\ E\exp(tV(\boldsymbol{\theta},b,\varepsilon)). \end{aligned}$$

Observing that (see the proof of (A.1) in Shao (1997))

$$\lim_{k\to\infty}\sup_{b\ge k}\sup_{||\boldsymbol{\theta}||=1}\inf_{0\le t\le m} E\exp(tV(\boldsymbol{\theta},b,\varepsilon))=0 \tag{2.13}$$

uniformly in $0\le\varepsilon\le 1/2$ and $m\ge 1$, we have

$$\limsup_{\varepsilon\downarrow 0}\sup_{b\ge 0}\sup_{||\boldsymbol{\theta}||=1}\inf_{0\le t\le m} E\exp(tV(\boldsymbol{\theta},b,\varepsilon))=\sup_{b\ge 0}\sup_{||\boldsymbol{\theta}||=1}\inf_{0\le t\le m} E\exp(tV(\boldsymbol{\theta},b,0)).$$

Finally by Lemma 4 of Chernoff (1952) and (**??**) again,

$$\lim_{m\to\infty}\sup_{b\ge 0}\sup_{||\boldsymbol{\theta}||=1}\inf_{0\le t\le m} E\exp(tV(\boldsymbol{\theta},b,0))=\sup_{b\ge 0}\sup_{||\boldsymbol{\theta}||=1}\inf_{0\le t} E\exp(tV(\boldsymbol{\theta},b,0))=K(\alpha).$$

This proves Theorem **??**. □

Proof of Theorem ??. By (**??**), it suffices to show that for all $y_n\to\infty$, $y_n=o(n)$,

$$\lim_{n\to\infty} y_n^{-1}\ln P\Big(\sup_{||\boldsymbol{\theta}||=1}\frac{\sum_{i=1}^n\boldsymbol{\theta}'\boldsymbol{X}_i}{(\sum_{i=1}^n(\boldsymbol{\theta}'\boldsymbol{X}_i)^2)^{1/2}}\ge y_n^{1/2}\Big)=-\frac{1}{2} \tag{2.14}$$

Recall that for any R^d-valued random variable $\boldsymbol{X}$

$$E||\boldsymbol{X}||^2 1\{||\boldsymbol{X}||\le x\}\text{ slowly varying }\Leftrightarrow\ x^2P(||\boldsymbol{X}||>x)/E||\boldsymbol{X}||^2 1\{||\boldsymbol{X}||\le x\}\to 0 \tag{2.15}$$

(see for example, Theorem 1.8.1 of Bingham et al. (1987)). Since $h(x)=E||\boldsymbol{X}||^2 1\{||\boldsymbol{X}||\le x\}$ is slowly varying, it follows from (**??**) and (**??**) that for every $\boldsymbol{\theta}\in R^d$ with $||\boldsymbol{\theta}||=1$,

$$\begin{aligned} x^2P(|\boldsymbol{\theta}'\boldsymbol{X}|>x) &\le x^2P(||\boldsymbol{X}||>x)=o(h(x))\\ &= o(E(\boldsymbol{\theta}'\boldsymbol{X})^2 1\{||\boldsymbol{X}||\le x\})=o(E(\boldsymbol{\theta}'\boldsymbol{X})^2 1\{|\boldsymbol{\theta}'\boldsymbol{X}|\le x\}).\end{aligned}$$

Applying (**??**) for the R-valued $\boldsymbol{\theta}'\boldsymbol{X}$, we see that $E(\boldsymbol{\theta}'\boldsymbol{X})^2 1\{|\boldsymbol{\theta}'\boldsymbol{X}|\le x\}$ is slowly varying. With $E\boldsymbol{\theta}'\boldsymbol{X}=0$ it follows from Theorem 3.1 of Shao (1997) that

$$\begin{aligned}&\liminf_{n\to\infty} y_n^{-1}\ln P\Big(\sup_{||\boldsymbol{\theta}||=1}\frac{\sum_{i=1}^n\boldsymbol{\theta}'\boldsymbol{X}_i}{(\sum_{i=1}^n(\boldsymbol{\theta}'\boldsymbol{X}_i)^2)^{1/2}}\ge y_n^{1/2}\Big)\\ &\ge\liminf_{n\to\infty} y_n^{-1}\ln P\Big(\frac{\sum_{i=1}^n\boldsymbol{\theta}'\boldsymbol{X}_i}{(\sum_{i=1}^n(\boldsymbol{\theta}'\boldsymbol{X}_i)^2)^{1/2}}\ge y_n^{1/2}\Big)=-\frac{1}{2}\,,\end{aligned}$$

establishing the lower bound in (**??**). Since $y_n=o(n)$ there exists $z_n\to\infty$ such that $y_n=(1+o(1))nz_n^{-2}h(z_n)$ (cf. Proposition 1.3.6 and Theorems 1.8.2, 1.8.5 of Bingham et al. (1987)). It thus suffices to prove the complementary upper bound in (**??**) for $y_n=nz_n^{-2}h(z_n)$ and any $z_n\to\infty$. Fixing $z_n\to\infty$ and $0<\varepsilon<1/4$ set

$$\xi_i(\boldsymbol{\theta}):=\xi_i(\boldsymbol{\theta},z_n)=\boldsymbol{\theta}'\boldsymbol{X}_i 1\{||\boldsymbol{X}_i||\le\varepsilon z_n\}\ .$$

Similarly to (**??**), we see that

$$P\Big(\sup_{||\boldsymbol{\theta}||=1}\frac{\sum_{i=1}^n \boldsymbol{\theta}'\boldsymbol{X}_i}{(\sum_{i=1}^n(\boldsymbol{\theta}'\boldsymbol{X}_i)^2)^{1/2}}\geq y_n^{1/2}\Big) \tag{2.16}$$

$$\begin{aligned}
&\leq P\Big(\sup_{||\boldsymbol{\theta}||=1}\Big\{\sum_{i=1}^n \xi_i(\boldsymbol{\theta})-(1-\varepsilon)y_n^{1/2}(\sum_{i=1}^n \xi_i^2(\boldsymbol{\theta}))^{1/2}\Big\}\geq 0\Big)\\
&+ P\Big(\sum_{i=1}^n 1\{||\boldsymbol{X}_i||>\varepsilon z_n\}\geq \varepsilon^2 y_n\Big)\\
&:= J_1+J_2
\end{aligned}$$

With $y_n = nz_n^{-2}h(z_n)$ and $z_n\to\infty$, it follows by (**??**) that

$$y_n^{-1}\ln J_2 \leq \varepsilon^2\ln\Big(3z_n^2P(||\boldsymbol{X}||>\varepsilon z_n)/(\varepsilon^2h(z_n))\Big)$$

With $h(x)$ slowly varying, it follows from (**??**) that $(\varepsilon z_n)^2P(||\boldsymbol{X}||\geq\varepsilon z_n)/h(z_n)\to 0$ as $n\to\infty$, hence

$$\limsup_{n\to\infty} y_n^{-1}\ln J_2 = -\infty\,. \tag{2.17}$$

Let $\eta\in(0,1/(4d))$. Consider a finite η-cover $\mathcal{H}$ of $\{\boldsymbol{\theta}:\ \boldsymbol{\theta}\in R^d,\ ||\boldsymbol{\theta}||=1\}$ with respect to the maximum norm in R^d. Thus, for any $\boldsymbol{\theta}\in R^d$ with $||\boldsymbol{\theta}||=1$, there exists $\boldsymbol{\theta}_0\in\mathcal{H}$ such that

$$||\boldsymbol{\theta}-\boldsymbol{\theta}_0||_\infty\leq\eta \text{ and } ||\boldsymbol{\theta}_0||=1\,.$$

Since $\sum_{i=1}^n\xi_i(\boldsymbol{\theta})$ is linear in $\boldsymbol{\theta}$, it follows that

$$\sup_{||\boldsymbol{\theta}-\boldsymbol{\theta}_0||_\infty\leq\eta}\sum_{i=1}^n\xi_i(\boldsymbol{\theta}) = \max_{\boldsymbol{\vartheta}\in\mathcal{H}(\boldsymbol{\theta}_0)}\sum_{i=1}^n\xi_i(\boldsymbol{\vartheta})\,,$$

where $\mathcal{H}(\boldsymbol{\theta}_0):=\{\boldsymbol{\theta}_0+\eta\boldsymbol{\delta}:\ \ \boldsymbol{\delta}\in\{-1,1\}^d\}$. Consequently,

$$P\Big(\sup_{||\boldsymbol{\theta}-\boldsymbol{\theta}_0||_\infty\leq\eta}\Big\{\sum_{i=1}^n\xi_i(\boldsymbol{\theta})-(1-\varepsilon)y_n^{1/2}(\sum_{i=1}^n\xi_i^2(\boldsymbol{\theta}))^{1/2}\Big\}\geq 0\Big) \tag{2.18}$$

$$\begin{aligned}
&\leq P\Big(\max_{\boldsymbol{\vartheta}\in\mathcal{H}(\boldsymbol{\theta}_0)}\sum_{i=1}^n\xi_i(\boldsymbol{\vartheta})\geq(1-\varepsilon)y_n^{1/2}\inf_{||\boldsymbol{\theta}-\boldsymbol{\theta}_0||_\infty\leq\eta}(\sum_{i=1}^n\xi_i^2(\boldsymbol{\theta}))^{1/2}\Big)\\
&\leq \sum_{\boldsymbol{\vartheta}\in\mathcal{H}(\boldsymbol{\theta}_0)}\Big\{P\Big(\sum_{i=1}^n\xi_i(\boldsymbol{\vartheta})\geq(1-\varepsilon)^2y_n^{1/2}(n\,E\xi^2(\boldsymbol{\vartheta}))^{1/2}\Big)\\
&\qquad +P\Big(\inf_{||\boldsymbol{\theta}-\boldsymbol{\vartheta}||_\infty\leq 2\eta}\sum_{i=1}^n\xi_i^2(\boldsymbol{\theta})\leq(1-\varepsilon)n\,E\xi^2(\boldsymbol{\vartheta})\Big)\Big\}\\
&:= \sum_{\boldsymbol{\vartheta}\in\mathcal{H}(\boldsymbol{\theta}_0)}\Big\{J_{1,1}(\boldsymbol{\vartheta})+J_{1,2}(\boldsymbol{\vartheta})\Big\}\,.
\end{aligned}$$

Recall that $E||\boldsymbol{X}||1\{||\boldsymbol{X}||>x\} = xP(||\boldsymbol{X}||>x)+\int_x^\infty P(||\boldsymbol{X}||>y)dy = o(h(x)/x)$ (cf. Proposition 1.5.10 of Bingham et al. (1987), or (4.5) of Shao (1997)). Thus, with $E\boldsymbol{X}=0$ it follows that

$$|E\xi(\boldsymbol{\vartheta})| = |E\boldsymbol{\vartheta}'\boldsymbol{X}1\{||\boldsymbol{X}||>\varepsilon z_n\}|\leq||\boldsymbol{\vartheta}||E||\boldsymbol{X}||1\{||\boldsymbol{X}||>\varepsilon z_n\} = o(h(\varepsilon z_n)/(\varepsilon z_n))$$

By assumption (??) we have $E\xi^2(\boldsymbol{\vartheta}) \geq c_0 h(\varepsilon z_n)/2$ and hence

$$\sum_{i=1}^n E\xi_i(\boldsymbol{\vartheta}) \leq \varepsilon(1-\varepsilon)^2 y_n^{1/2}(n\, E\xi^2(\boldsymbol{\vartheta}))^{1/2}\,,$$

for all n large enough and all $\boldsymbol{\vartheta} \in \mathcal{H}(\boldsymbol{\theta}_0)$, $\boldsymbol{\theta}_0 \in \mathcal{H}$. As $||\boldsymbol{\vartheta}|| \leq 1 + 1/(4\sqrt{d}) \leq 5/4$, $|\xi(\boldsymbol{\vartheta})| \leq (5/4)\varepsilon z_n$. It follows by (??) and Bernstein's inequality that for some $C < \infty$ and all n large enough, $\boldsymbol{\vartheta} \in \mathcal{H}(\boldsymbol{\theta}_0)$, $\boldsymbol{\theta}_0 \in \mathcal{H}$,

$$\begin{aligned} J_{1,1}(\boldsymbol{\vartheta}) &\leq P\Big(\sum_{i=1}^n (\xi_i(\boldsymbol{\vartheta}) - E\xi_i(\boldsymbol{\vartheta})) \geq (1-\varepsilon)^3 y_n^{1/2}(n\, E\xi^2(\boldsymbol{\vartheta}))^{1/2}\Big) \\ &\leq \exp\Big(-\frac{(1-\varepsilon)^6 y_n\, n\, E\xi^2(\boldsymbol{\vartheta})}{2n\, E\xi^2(\boldsymbol{\vartheta}) + 2(1-\varepsilon)^3(y_n\, n\, E\xi^2(\boldsymbol{\vartheta}))^{1/2}(\varepsilon z_n)}\Big) \\ &\leq \exp\Big(-\frac{(1-\varepsilon)^6 y_n}{2(1+C\varepsilon)}\Big). \end{aligned} \tag{2.19}$$

As to $J_{1,2}(\boldsymbol{\vartheta})$, noting that

$$\inf_{||\boldsymbol{\theta}-\boldsymbol{\vartheta}||_\infty \leq 2\eta} \sum_{i=1}^n \xi_i^2(\boldsymbol{\theta}) \geq \sum_{i=1}^n \xi_i^2(\boldsymbol{\vartheta}) - 8\sqrt{d}\eta \sum_{i=1}^n ||\boldsymbol{X}_i||^2 1\{||\boldsymbol{X}_i|| \leq \varepsilon z_n\},$$

we have

$$\begin{aligned} J_{1,2}(\boldsymbol{\vartheta}) &\leq P\Big(\sum_{i=1}^n \xi_i^2(\boldsymbol{\vartheta}) \leq (1-\varepsilon/2) n\, E\xi^2(\boldsymbol{\vartheta})\Big) \\ &+ P\Big(8\sqrt{d}\eta \sum_{i=1}^n ||\boldsymbol{X}_i||^2 1\{||\boldsymbol{X}_i|| \leq \varepsilon z_n\} \geq \varepsilon n\, E\xi^2(\boldsymbol{\vartheta})/2\Big)\,. \end{aligned} \tag{2.20}$$

Recall that

$$E\xi^4(\boldsymbol{\vartheta}) \leq ||\boldsymbol{\vartheta}||^4 E||\boldsymbol{X}||^4 1\{||\boldsymbol{X}|| \leq \varepsilon z_n\} = o((\varepsilon z_n)^2 h(z_n)) \tag{2.21}$$

(cf. Proposition 1.5.10 of Bingham et al. (1987)). Using (??), (??) and Bernstein's inequality, we see that for all sufficiently large n, $\boldsymbol{\vartheta} \in \mathcal{H}(\boldsymbol{\theta}_0)$, $\boldsymbol{\theta}_0 \in \mathcal{H}$,

$$\begin{aligned} &P\Big(\sum_{i=1}^n \xi_i^2(\boldsymbol{\vartheta}) \leq (1-\varepsilon/2) n\, E\xi^2(\boldsymbol{\vartheta})\Big) \\ &\leq \exp\Big(-\frac{(\varepsilon n\, E\xi^2(\boldsymbol{\vartheta})/2)^2}{2nE\xi^4(\boldsymbol{\vartheta}) + \varepsilon n\, E\xi^2(\boldsymbol{\vartheta})(\varepsilon z_n)^2}\Big) \\ &\leq \exp\Big(-\frac{(nE\xi^2(\boldsymbol{\vartheta}))^2}{o(1) n z_n^2 h(z_n)}\Big) + \exp\Big(-\frac{n\, E\xi^2(\boldsymbol{\vartheta})}{4\varepsilon z_n^2}\Big) \\ &\leq \exp\Big(-y_n c_0^2/o(1)\Big) + \exp\Big(-y_n c_0/(8\varepsilon)\Big)\,. \end{aligned} \tag{2.22}$$

Similarly, for η sufficiently small, say $\eta < \varepsilon c_0/(32\sqrt{d})$, by (**??**),

$$\begin{aligned} & P\Big(\sum_{i=1}^{n} ||\boldsymbol{X}_i||^2 1\{||\boldsymbol{X}_i|| \le \varepsilon z_n\} \ge \frac{\varepsilon n\, E\xi^2(\boldsymbol{\vartheta})}{16\sqrt{d}\eta}\Big) \\ \le\ & P\Big(\sum_{i=1}^{n} \big(||\boldsymbol{X}_i||^2 1\{||\boldsymbol{X}_i|| \le \varepsilon z_n\} - E||\boldsymbol{X}_i||^2 1\{||\boldsymbol{X}_i|| \le \varepsilon z_n\}\big) \ge nh(\varepsilon z_n)\Big) \\ \le\ & \exp\Big(- y_n/(2\varepsilon^2 + o(1))\Big) \end{aligned} \quad (2.23)$$

Combining (**??**), (**??**), (**??**), (**??**) and (**??**) yields for all ε small enough and n large enough,

$$J_1 = O(1) \exp\Big(- (1-\varepsilon)^6 y_n/(2(1 + C\varepsilon))\Big) \quad (2.24)$$

Taking $n \to \infty$ then $\varepsilon \to 0$ this proves the upper bound of (**??**). $\square$.

Proof of Theorem ??. By using the Ottaviani maximum inequality and following the proof of Theorem **??**, one can have a stronger version of (**??**): for arbitrary $0 < \varepsilon < 1/2$, there exist $0 < \delta < 1$, $y_0 > 1$ and n_0 such that for any $n \ge n_0$ and $y_0 < y < \delta n$,

$$P\Big(\sup_{n \le k \le (1+\delta)n} \sup_{||\boldsymbol{\theta}||=1} \frac{\sum_{i=1}^{k} \boldsymbol{\theta}' \boldsymbol{X}_i}{(\sum_{i=1}^{k} (\boldsymbol{\theta}' \boldsymbol{X}_i)^2)^{1/2}} \ge y^{1/2}\Big) \le \exp\Big(- (1-\varepsilon)y/2\Big). \quad (2.25)$$

Using the subsequence method it follows from (**??**) and the Borel-Cantelli lemma that

$$\limsup_{n\to\infty} \frac{T_n^2}{2 \log\log n} \le 1 \quad a.s.$$

As to the lower bound, it follows from the representation (**??**) and the self-normalized law of the iterated logarithm for $d = 1$ (see Theorem 1 of Griffin and Kuelbs (1989)). For a similar proof, see that of Corollary 5.2 of Dembo and Shao (1998).

Acknowledgements. The authors would like to thank two referees and the editor for their valuable comments.

References

[1] Anderson, T.W. (1984). *An introduction to Multivariate Analysis* (2nd ed.). Wiley, New York.

[2] Bercu, B., Gassiat, E. and Rio, E. (2002). Concentration inequalities, large and moderate deviations for self-normalized empirical processes. *Ann. Probab.* **30**, 1576–1604.

[3] Bingham, N. H., Goldie, C. M. and Teugels, J. L. (1987). Regular variation. Cambridge University Press, Cambridge.

[4] Chernoff, H. (1952). A measure of asymptotic efficiency for tests of a hypothesis based on the sum of obersevations. *Ann. Math. Stat.* **23**, 493-507.

[5] Chistyakov, G.P. and Götze, F. (2004a). On bounds for moderate deviations for Student's statistic. *Theory Probab. Appl.* **48**, 528-535.

[6] Chistyakov, G.P. and Götze, F. (2004b). Limit distributions of Studentized means. *Ann. Probab.* **32** (2004), 28-77.

[7] Dembo, A. and Shao, Q. M. (1998a). Self-normalized moderate deviations and lils. *Stoch. Proc. and Appl.* **75**, 51-65.

[8] Dembo, A. and Shao, Q. M. (1998b). Self-normalized large deviations in vector spaces. In: *Progress in Probability* (Eberlein, Hahn, Talagrand, eds) Vol. **43**, 27-32.

[9] Faure, M. (2002). Self-normalized large deviations for Markov chains. *Electronic J. Probab.* **7**, 1-31.

[10] Fujikoshi, Y. (1997). An asymptotic expansion for the distribution of Hotelling's T^2-statistic under nonnormality. *J. Multivariate Anal.* **61**, 187-193.

[11] Griffin, P. and Kuelbs, J. (1989). Self-normalized laws of the iterated logarithm. *Ann. Probab.* **17**, 1571–1601.

[12] Hahn, M.G. and Klass, M.J. (1980). Matrix normalization of sums of random vectors in the domain of attraction of the multivariate normal. *Ann. Probab.* **8**, 262-280.

[13] He, X. and Shao, Q. M. (1996). Bahadur efficiency and robustness of studentized score tests. *Ann. Inst. Statist. Math.* **48**, 295-314.

[14] Jing, B.Y., Shao, Q.M. and Wang, Q.Y. (2003). Self-normalized Cramér type large deviations for independent random variables. *Ann. Probab.* **31** , 2167-2215.

[15] Jing, B.Y., Shao, Q.M. and Zhou, W. (2004). Saddlepoint approximation for Student's t-statistic with no moment conditions. *Ann. Statist.* **32**, 2679-2711.

[16] Kano, Y. (1995). An asymptotic expansion of the distribution of Hotelling's T^2-statistic under general condition. *Amer. J. Math. Manage. Sci.* **15**, 317-341.

[17] Kariya, T. (1981). A robustness property of Hotelling's T^2-test. *Ann. Statist.* **9**, 210-213.

[18] Kiefer, J. and Schwartz, R. (1965). Admissible Bayes character of T^2- and R^2- and other fully invariant tests for classical normal prolems. *Ann. Math. Statist.* **36**, 747-760.

[19] Muirhead, R.J. (1982). *Aspects of Multivariate Statistical Theory.* John Wiley, New York.

[20] Robinson, J. and Wang, Q.Y. (2005). On the self-normalized Cramér-type large deviation. *J. Theoretic Probab.* **18**, 891-909.

[21] Sepanski, S. (1994). Asymptotics for Multivariate t-statistic and Hotelling's T^2-statistic under infinite second moments via bootstrapping. *J. Multivariate Anal.* **49**, 41-54.

[22] Shao, Q. M. (1997). Self-normalized large deviations. *Ann. Probab.* **25**, 285–328.

[23] Shao, Q. M. (1998). Recent developments in self-normalized limit theorems. In *Asymptotic Methods in Probability and Statistics* (editor B. Szyszkowicz), pp. 467 - 480. Elsevier Science.

[24] Shao, Q. M. (2004). Recent progress on self-normalized limit theorems. In *Probability, finance and insurance* (editors Tze Leung Lai, Hailiang Yang and Siu Pang Yung), pp. 50–68, World Sci. Publ., River Edge, NJ, 2004.

[25] Simaika, J.B. (1941). On an optimal property of two important statistical tests. *Biometrika* **32**, 70-80.

[26] Stein, C. (1956). The admissibility of Hotelling's T^2-test. *Ann. Math. Statist.* **27**, 616-623.

[27] Wang, Q.Y. (2005). Limit theorems for self-normalized large deviation. *Electronic J. Probab.* **10**, 1260-1285.