ORIGINAL PAPER

On maximum likelihood estimation of the concentration parameter of von Mises–Fisher distributions

Kurt Hornik · Bettina Grün

Abstract Maximum likelihood estimation of the concentration parameter of von Mises–Fisher distributions involves inverting the ratio $R_\nu = I_{\nu+1}/I_\nu$ of modified Bessel functions and computational methods are required to invert these functions using approximative or iterative algorithms. In this paper we use Amos-type bounds for R_ν to deduce sharper bounds for the inverse function, determine the approximation error of these bounds, and use these to propose a new approximation for which the error tends to zero when the inverse of R_ν is evaluated at values tending to 1 (from the left). We show that previously introduced rational bounds for R_ν which are invertible using quadratic equations cannot be used to improve these bounds.

1 Introduction

A random unit length vector in $\mathbb{R}^d$ has a *von Mises–Fisher* (or Langevin) distribution with parameter $\theta \in \mathbb{R}^d$ if its density with respect to the uniform distribution on the unit hypersphere $\mathbb{S}^{d-1} = \{x \in \mathbb{R}^d : \|x\| = 1\}$ is given by

K. Hornik
Institute for Statistics and Mathematics, WU Wirtschaftsuniversität Wien,
Vienna, Austria
e-mail: Kurt.Hornik@wu.ac.at

B. Grün (✉)
Institut für Angewandte Statistik, Johannes Kepler Universität Linz,
Linz, Austria
e-mail: Bettina.Gruen@jku.at

$$f(x|\theta) = e^{\theta'x} /_0F_1(; d/2; \|\theta\|^2/4),$$

where

$$_0F_1(; \nu; z) = \sum_{n=0}^{\infty} \frac{\Gamma(\nu)}{\Gamma(\nu+n)} \frac{z^n}{n!}$$

is a generalized hypergeometric series and related to the modified Bessel function of the first kind I_ν via

$$_0F_1(; \nu+1; \kappa^2/4) = \frac{I_\nu(\kappa)\Gamma(\nu+1)}{(\kappa/2)^\nu}$$

(e.g., Mardia and Jupp 1999, p. 168).

We note that the von Mises–Fisher distribution is commonly parametrized as $\theta = \kappa\mu$, where $\kappa = \|\theta\|$ and $\mu \in \mathbb{S}^{d-1}$ are the *concentration* and *mean direction* parameters, respectively (if $\theta \neq 0$, μ is uniquely determined as $\theta/\|\theta\|$).

Using the common parametrization by κ and μ, the log-likelihood of a sample $x_1, \ldots, x_n$ from the von Mises–Fisher distribution is given by

$$-n \log(_0F_1(; d/2; \kappa^2/4)) + \kappa\mu'r,$$

where $r = \sum_{i=1}^{n} x_i$ is the resultant vector (sum) of the x_i. Using recursions for the modified Bessel functions (e.g., Watson 1995, p. 71), one can show that the negative log-derivative of $\kappa \mapsto \log(_0F_1(; d/2; \kappa^2/4))$ equals $A_d(\kappa) = R_{d/2-1}(\kappa)$, where $R_\nu(t) = I_{\nu+1}(t)/I_\nu(t)$. The maximum likelihood estimators are thus obtained by taking $\hat{\mu} = r/\|r\|$ and solving

$$R_{d/2-1}(\hat{\kappa}) = \rho,$$

where $\rho = \|r\|/n$ is the mean resultant length (Schou 1978).

It can be shown (e.g., Schou 1978) that for $\nu \geq 0$, R_ν is strictly increasing, and satisfies the Riccati equation $R'_\nu(t) = 1 - ((2\nu+1)/t)R_\nu(t) - R_\nu(t)^2$. As R_ν and hence also its derivatives can efficiently be computed via its Perron or Gauss continued fraction representation (Gautschi and Slavik 1978; Tretter and Walster 1980; Song et al. 2012), solving $R_\nu(t) = \rho$ should conveniently be achievable by standard iterative root finding techniques, provided that good starting approximations are available (which is particularly important in the right tail of R_ν where R_ν is rather "flat").

Dhillon and Sra (2003) and subsequently Banerjee et al. (2005) suggest the approximation

$$R_\nu^{-1}(\rho) \approx \frac{\rho}{1-\rho^2}(2(\nu+1) - \rho^2) =: Q_\nu(\rho) \tag{1}$$

obtained by truncating the Gauss continued fraction representation of R_ν and adding a correction term "determined empirically", pointing out that this initial approximation can subsequently be improved by Newton–Raphson iterations. Sra (2012) suggests to use exactly two such iterations.

Tanabe et al. (2007) show that for $v \geq 0$ and $0 \leq \rho < 1$,

$$R_v^{-1}(\rho) = \frac{\rho}{1 - \rho^2}(2(v+1) - c)$$

for some suitable $0 \leq c = c(v, \rho) \leq 2$, or equivalently,

$$\frac{2v\rho}{1 - \rho^2} \leq R_v^{-1}(\rho) \leq \frac{2(v+1)\rho}{1 - \rho^2}, \tag{2}$$

with the Dhillon–Sra approximation assuming $c \approx \rho^2$. The upper and lower bound differ by $2\rho/(1 - \rho^2)$ which is independent of v but tends to infinity as $\rho \to 1-$. Tanabe et al. (2007) also suggest to use the "mid-point" approximation with $c = 1$, i.e., $R_v^{-1}(\rho) \approx (2v + 1)\rho/(1 - \rho^2)$ as the starting value for iterative schemes for solving $R_v(t) = \rho$, such as the fixed-point iteration $t_{n+1} = t_n \rho / R_v(t_n)$.

In this paper, we use a family of bounds for R_v first introduced in Amos (1974) to provide substantially sharper bounds for R_v^{-1}, which have approximation error at most $3\rho/2$, and use these results to suggest a new approximation. We establish that these improved bounds also hold for the Dhillon–Sra approximation, which thus has the same maximal approximation error. We also show that the error of the suggested new approximation tends to zero for $\rho \to 1-$, whereas the error tends to $-1/2$ for the Dhillon–Sra approximation, which thus is too large for large ρ. Finally, we investigate whether the rational bounds for R_v developed by Nåsell (1978) can be used to obtain improved explicit bounds for R_v^{-1}, and show that for the rational bounds which can be inverted by solving quadratic equations, no such improvement is possible.

2 Amos-type bounds

Let

$$G_{\alpha,\beta}(t) = \frac{t}{\alpha + \sqrt{t^2 + \beta^2}},$$

where in what follows we take $\beta \geq 0$ without loss of generality. Amos (1974) gives the bounds

$$G_{v+1/2, v+3/2}(\kappa) \leq R_v(\kappa) \leq G_{v+1/2, v+1/2}(\kappa), \qquad \kappa, v \geq 0$$

(Equation 16) and

$$G_{v+1, v+1}(\kappa) \leq R_v(\kappa) \leq G_{v, v+2}(\kappa) \leq G_{v,v}(\kappa), \qquad \kappa, v \geq 0$$

(Equations 9 and 11). The bounds are actually valid for larger v domains (see for example Nåsell 1978; Yuan and Kalbfleisch 2000). It is trivial that $G_{v+1/2, v+1/2}(t) < G_{v,v}(t)$ for all $t > 0$. For $\Delta(t) = (v + 1/2) + \sqrt{t^2 + (v + 3/2)^2} - ((v + 1) + \sqrt{t^2 + (v + 1)^2})$ we have $\Delta(0) = 0$ and $\Delta'(t) = t/\sqrt{t^2 + (v + 3/2)^2} -$

$t/\sqrt{t^2 + (v+1)^2}$, which is negative for all $t > 0$. Thus, $\Delta(t) < 0$ for all $t > 0$ and hence $G_{v+1/2,v+3/2}(t) > G_{v+1,v+1}(t)$ for all $t > 0$.

Let

$$\beta_{SS}(v) = \sqrt{(v + 1/2)(v + 3/2)}.$$

Simpson and Spector (1984) show that with $v_v(t) = t/R_v(t)$, $v_v(t)^2 - (2v+1)v_v(t) - (t^2 + v + 1/2) > 0$ for all $v \geq 0$ and $t > 0$, which is readily seen to imply $v_v(t) \geq v + 1/2 + \sqrt{t^2 + \beta_{SS}(v)^2}$ and hence $R_v(t) \leq G_{v+1/2,\beta_{SS}(v)}(t)$, which is clearly smaller than $G_{v+1/2,v+1/2}(t)$ for all $t > 0$.

Altogether, we thus have that for $v \geq 0$ and $t \geq 0$,

$$G_{v+1/2,v+3/2}(t) \leq R_v(t) \leq \min\left(G_{v,v+2}(t), G_{v+1/2,\beta_{SS}(v)}(t)\right). \tag{3}$$

What makes these Amos-type bounds particularly attractive is that they can be inverted explicitly, as shown in the following lemma.

Lemma 1 *Let $\alpha \geq 0$ and $\alpha + \beta > 0$. Then $G_{\alpha,\beta}$ is strictly increasing on $[0, \infty)$, and for all $0 \leq \rho < 1$ the equation $G_{\alpha,\beta}(t) = \rho$ has a unique solution $t = G_{\alpha,\beta}^{-1}(\rho)$ given by*

$$G_{\alpha,\beta}^{-1}(\rho) = \frac{\rho}{1 - \rho^2}\left(\alpha + \sqrt{\rho^2\alpha^2 + (1 - \rho^2)\beta^2}\right). \tag{4}$$

Proof The derivative of $G_{\alpha,\beta}$ is given by

$$\begin{aligned}
G_{\alpha,\beta}'(t) &= \frac{1}{\alpha + \sqrt{t^2 + \beta^2}} - \frac{t}{(\alpha + \sqrt{t^2 + \beta^2})^2} \frac{2t}{2\sqrt{t^2 + \beta^2}} \\
&= \frac{\alpha\sqrt{t^2 + \beta^2} + \beta^2}{(\alpha + \sqrt{t^2 + \beta^2})^2\sqrt{t^2 + \beta^2}},
\end{aligned}$$

where the numerator has value $\beta(\alpha + \beta)$ at $t = 0$ and hence is positive for all $t > 0$ if and only if $\alpha \geq 0$ and $\alpha + \beta > 0$ in which case $G_{\alpha,\beta}$ is strictly increasing, and as $G_{\alpha,\beta}(0) = 0$ and $\lim_{t\to\infty} G_{\alpha,\beta}(t) = 1$ the equation $G_{\alpha,\beta}(t) = \rho$ has a unique solution for $0 \leq \rho < 1$. Now $G_{\alpha,\beta}(t) = \rho$ iff $t = \rho(\alpha + \sqrt{t^2 + \beta^2})$, giving t as the larger root of the quadratic equation $(1 - \rho^2)t^2 - 2\rho\alpha t + \rho^2(\alpha^2 - \beta^2) = 0$, so that

$$\begin{aligned}
G_{\alpha,\beta}^{-1}(\rho) &= \frac{1}{2(1 - \rho^2)}\left(2\rho\alpha + \sqrt{4\rho^2\alpha^2 - 4(1 - \rho^2)\rho^2(\alpha^2 - \beta^2)}\right) \\
&= \frac{\rho}{1 - \rho^2}\left(\alpha + \sqrt{\rho^2\alpha^2 + (1 - \rho^2)\beta^2}\right)
\end{aligned}$$

(the smaller root does not converge to ∞ as $\rho \to 1-$).

Theorem 1 *Let $v \geq 0$ and $0 \leq \rho < 1$. Then*

$$\max\left(G_{v,v+2}^{-1}(\rho), G_{v+1/2,\beta_{SS}(v)}^{-1}(\rho)\right) \leq R_v^{-1}(\rho) \leq G_{v+1/2,v+3/2}^{-1}(\rho). \tag{5}$$

Proof This follows immediately from combining the previous lemma with Eq. 3.

The above result substantially improves the results of Tanabe et al. (2007). Using Eq. 4, $G_{\alpha,\alpha}^{-1}(\rho) = 2\alpha\rho/(1 - \rho^2)$, so that the lower and upper bound in Eq. 2 equal $G_{\nu,\nu}^{-1}(\rho)$ and $G_{\nu+1,\nu+1}^{-1}(\rho)$, respectively, and hence correspond to $G_{\nu+1,\nu+1}(t) \leq R_\nu(t) \leq G_{\nu,\nu}(t)$, which was already shown to be strictly weaker than the bounds in Eq. 3. We also see that the "mid-point approximation" $R_\nu^{-1}(\rho) \approx (2\nu + 1)\rho/(1 - \rho^2)$ equals $G_{\nu+1/2,\nu+1/2}^{-1}(\rho)$, which for positive ρ is strictly smaller than $G_{\nu+1/2,\beta_{SS}(\nu)}^{-1}(\rho)$, and hence strictly under-estimates $R_\nu^{-1}(\rho)$.

Let

$$g(\nu) = \frac{(\nu + 3/2)}{2\nu + 1}.$$

Then g is monotonically decreasing on $[0, \infty)$ with $g(0) = 3/2$ and $\lim_{\nu\to\infty} g(\nu) = 1/2$.

Theorem 2 *Let* $\nu \geq 0$ *and* $0 \leq \rho < 1$. *Then*

$$0 \leq G_{\nu+1/2,\nu+3/2}^{-1}(\rho) - G_{\nu+1/2,\beta_{SS}(\nu)}^{-1}(\rho) \leq \rho g(\nu)$$

and for $\beta_{SS}(\nu) \leq \beta \leq \nu + 3/2$,

$$\left| R_\nu^{-1}(\rho) - G_{\nu+1/2,\beta}^{-1}(\rho) \right| \leq \rho g(\nu).$$

Proof Using the mean value theorem, for $u_1 \geq u_0 > 0$ and $\alpha \geq 0$ with a suitable $\tilde{u} \in (u_0, u_1)$,

$$0 \leq \sqrt{\alpha + u}\Big|_{u=u_0}^{u=u_1} = \frac{u_1 - u_0}{2\sqrt{\alpha + \tilde{u}}} \leq \frac{u_1 - u_0}{2\sqrt{\alpha + u_0}}$$

and hence

$$\begin{aligned}
0 &\leq G_{\nu+1/2,\nu+3/2}^{-1}(\rho) - G_{\nu+1/2,\beta_{SS}(\nu)}^{-1}(\rho) \\
&= \frac{\rho}{1 - \rho^2} \sqrt{(\nu + 1/2)^2\rho^2 + \beta^2(1 - \rho^2)}\,\Big|_{\beta=\beta_{SS}(\nu)}^{\beta=\nu+3/2} \\
&\leq \frac{\rho}{1 - \rho^2} \frac{(1 - \rho^2)((\nu + 3/2)^2 - (\nu + 1/2)(\nu + 3/2))}{2\sqrt{(\nu + 1/2)^2\rho^2 + \tilde{\beta}^2(1 - \rho^2)}} \\
&\leq \frac{\rho(\nu + 3/2)}{2(\nu + 1/2)}.
\end{aligned}$$

For $\beta_{SS}(\nu) \leq \beta \leq \nu + 3/2$, both R_ν^{-1} and $G_{\nu+1/2,\beta}^{-1}$ are bounded below by $G_{\nu+1/2,\beta_{SS}(\nu)}^{-1}$ and above by $G_{\nu+1/2,\nu+3/2}^{-1}$, implying that $|R_\nu^{-1} - G_{\nu+1/2,\beta}^{-1}| \leq G_{\nu+1/2,\nu+3/2}^{-1} - G_{\nu+1/2,\beta_{SS}(\nu)}^{-1}$, whence the result from the bounds on this difference.

Corollary 1 *Let $v \geq 0$ and $0 \leq \rho < 1$. Then*

$$0 \leq R_v^{-1}(\rho) - \max\left(G_{v,v+2}^{-1}(\rho), G_{v+1/2, \beta_{SS}(v)}^{-1}(\rho)\right) \leq \rho g(v).$$

Proof Immediate from Theorems 1 and 2.

Theorem 3 *Let $v \geq 0$ and $0 \leq \rho < 1$. Then*

$$\max\left(G_{v,v+2}^{-1}(\rho), G_{v+1/2, \beta_{SS}(v)}^{-1}(\rho)\right) \leq Q_v(\rho) \leq G_{v+1/2, v+3/2}^{-1}(\rho).$$

Proof For $\beta_\alpha(\rho) = \sqrt{(\alpha+1)^2 - \rho^2}$,

$$\alpha^2 \rho^2 + \beta_\alpha(\rho)^2(1 - \rho^2) = \alpha^2 \rho^2 + ((\alpha+1)^2 - \rho^2)(1 - \rho^2) = (\alpha + 1 - \rho^2)^2.$$

Hence, $\alpha + \sqrt{\alpha^2 \rho^2 + \beta_\alpha(\rho)^2(1 - \rho^2)} = 2\alpha + 1 - \rho^2$ so that in particular,

$$Q_v(\rho) = G_{v+1/2, \beta_{v+1/2}(\rho)}^{-1}(\rho).$$

As clearly $\beta_{SS}(v) \leq \beta_{v+1/2}(\rho) \leq v + 3/2$ for all $0 \leq \rho \leq 1$, we thus obtain $G_{v+1/2, \beta_{SS}(v)}^{-1}(\rho) \leq G_{v+1/2, \beta_{v+1/2}(\rho)}^{-1}(\rho) = Q_v(\rho) \leq G_{v+1/2, v+3/2}^{-1}(\rho)$.

Writing $\Delta(\sigma) = \sigma + \sqrt{(v+2)^2 - 4(v+1)\sigma} - (v+2)$ we have $G_{v,v+2}^{-1}(\rho) - Q_v(\rho) = \Delta(\rho^2)\rho/(1 - \rho^2)$. As $\Delta'(\sigma) = 1 - 2(v+1)/\sqrt{(v+2)^2 - 4(v+1)\sigma}$ and $\Delta''(\sigma) = -4(v+1)^2((v+2)^2 - 4(v+1)\sigma)^{-3/2}$, Δ is strictly concave with its unique maximum at the solution σ^* of $\Delta'(\sigma) = 0$, or equivalently $(v+2)^2 - 4(v+1)\sigma = 4(v+1)^2$, from which

$$\sigma^* = \frac{(v+2)^2 - 4(v+1)^2}{4(v+1)} = \frac{-3v^2 - 4v}{4(v+1)},$$

which is non-positive for $v \geq 0$. Thus, Δ is decreasing on $[0, 1]$. As $\Delta(0) = 0$, we obtain that for $0 < \rho < 1$, $\Delta(\rho^2) < 0$ and hence $G_{v,v+2}^{-1}(\rho) < Q_v(\rho)$.

Corollary 2 *Let $v \geq 0$ and $0 \leq \rho < 1$. Then*

$$\left| R_v^{-1}(\rho) - Q_v(\rho) \right| \leq \rho g(v).$$

Proof Straightforward from combining Theorems 1, 2 and 3.

We thus see that the Dhillon–Sra approximation is not invalidated by the available inverse Amos-type bounds (in the sense of being outside the range provided by these bounds), and has the same maximal approximation error as these bounds (indicating that it is indeed a good approximation).

Theorem 4 *Let $v \geq 0$. Then as $\rho \to 1-$,*

$$R_v^{-1}(\rho) - G_{v+1/2, \beta_{SS}(v)}^{-1}(\rho) = O(\rho - 1)$$

and

$$R_v^{-1}(\rho) - Q_v(\rho) = -\frac{1}{2} + O(\rho - 1).$$

Proof Using the asymptotic expansion of I_v for large argument (e.g.,Watson 1995, Formula 7.23.2), one can show that for arbitrary v,

$$R_v(t) = 1 - \frac{v + 1/2}{t} + \frac{v^2 - 1/4}{2t^2} + O(1/t^3), \qquad t \to \infty, \tag{6}$$

see also Schou (1978, Eq. 6, assuming $v \geq 0$). Thus, we have $R_v^{-1}(\rho) = \omega_{-1}/(\rho-1)+$ $\omega_0 + O(\rho-1)$ as $\rho \to 1-$ with $\omega_{-1} \neq 0$, and the coefficients of this approximation can be determined by rewriting $\rho = R_v(t) \approx 1 + \alpha_1/t + \alpha_2/t^2$ as $(\rho-1)t^2 - \alpha_1 t - \alpha_2 \approx 0$ to obtain

$$t = R_v^{-1}(\rho) \approx \frac{\alpha_1 + \sqrt{\alpha_1^2 + 4\alpha_2(\rho - 1)}}{2(\rho - 1)} = \frac{\alpha_1}{\rho - 1} + \frac{\alpha_2}{\alpha_1} + O(\rho - 1),$$

giving (with $\alpha_1 = -(v + 1/2)$ and $\alpha_2 = (v - 1/2)(v + 1/2)/2$)

$$R_v^{-1}(\rho) = -\frac{v + 1/2}{\rho - 1} - \frac{v - 1/2}{2} + O(\rho - 1), \qquad \rho \to 1-.$$

For $\rho \to 1-$,

$$\frac{\rho}{1 - \rho^2} = -\frac{(\rho - 1) + 1}{(\rho - 1)(2 + (\rho - 1))}$$

$$= -\frac{1}{2}\left(\frac{1}{\rho - 1} + 1\right)\left(1 - \frac{\rho - 1}{2} + \frac{(\rho - 1)^2}{4} + O((\rho - 1)^3)\right)$$

$$= -\frac{1}{2}\left(\frac{1}{\rho - 1} + \frac{1}{2} - \frac{\rho - 1}{4} + O((\rho - 1)^2)\right).$$

Hence, if $f(\rho) = \delta_0 + \delta_1(\rho - 1) + O((\rho - 1)^2)$ as $\rho \to 1$,

$$\frac{\rho}{1 - \rho^2} f(\rho) = -\frac{\delta_0}{2}\frac{1}{\rho - 1} - \frac{\delta_0 + 2\delta_1}{4} + O(\rho - 1)$$

as $\rho \to 1-$. In particular, as for $\alpha > 0$

$$
\begin{aligned}
\sqrt{\rho^2 \alpha^2 + (1 - \rho^2)\beta^2} &= \sqrt{\alpha^2 + (\beta^2 - \alpha^2)(1 - \rho^2)} \\
&= \alpha \left(1 + \frac{\beta^2 - \alpha^2}{2\alpha^2}(1 - \rho^2) + O((\rho - 1)^2) \right) \\
&= \alpha - \frac{\beta^2 - \alpha^2}{\alpha}(\rho - 1) + O((\rho - 1)^2)
\end{aligned}
$$

as $\rho \to 1$,

$$
\begin{aligned}
G_{\alpha,\beta}^{-1}(\rho) &= \frac{\rho}{1 - \rho^2} \left(2\alpha - \frac{\beta^2 - \alpha^2}{\alpha}(\rho - 1) + O((\rho - 1)^2) \right) \\
&= -\frac{\alpha}{\rho - 1} - \frac{1}{4} \left(2\alpha - 2\frac{\beta^2 - \alpha^2}{2\alpha} \right) + O(\rho - 1) \\
&= -\frac{\alpha}{\rho - 1} - \alpha + \frac{\beta^2}{2\alpha} + O(\rho - 1), \qquad \rho \to 1-.
\end{aligned}
$$

For $\alpha = \nu + 1/2$ and $\beta = \beta_{SS}(\nu)$ we have $-\alpha + \beta^2/(2\alpha) = -\nu/2 + 1/4$ so that

$$
G_{\nu+1/2,\beta_{SS}(\nu)}^{-1}(\rho) = -\frac{\nu + 1/2}{\rho - 1} - \frac{\nu - 1/2}{2} + O(\rho - 1)
$$

and hence $R_\nu^{-1}(\rho) - G_{\nu+1/2,\beta_{SS}(\nu)}^{-1}(\rho) = O(\rho - 1)$ as $\rho \to 1-$.

As $2(\nu + 1) - \rho^2 = (2\nu + 1) - 2(\rho - 1) - (\rho - 1)^2$,

$$
Q_\nu(\rho) = -\frac{\nu + 1/2}{\rho - 1} - \frac{2\nu - 3}{4} + O(\rho - 1), \qquad \rho \to 1-.
$$

As

$$
-\frac{\nu - 1/2}{2} - \left(-\frac{2\nu - 3}{4} \right) = -\frac{1}{2},
$$

we thus have $R_\nu^{-1}(\rho) - Q_\nu(\rho) = -1/2 + O(\rho - 1)$ as $\rho \to 1-$, and the proof is complete.

3 Nåsell bounds

Nåsell (1978) gives families $L_{\nu,k,m}$ and $U_{\nu,k,m}$ of *rational* lower and upper bounds for R_ν, which converge to R_ν as $m \to \infty$ or $k \to \infty$ (Nåsell 1978, Theorems 2 and 3). These bounds can be used for obtaining bounds for R_ν^{-1} by applying numerical root finding techniques either directly to equations of the form $P(t)/Q(t) = \rho$ with polynomials P and Q, or by rewriting the equations of this form as $R(t) = P(t) -$

$\rho Q(t) = 0$ and then determining a suitable root of the polynomial R. "Simple" closed form expressions can be obtained when root finding amounts to solving a quadratic equation.

As $R_\nu(t)$ tends to 0 and 1 for $t \to 0$ and ∞, respectively, we thus restrict ourselves to Nåsell bounds exhibiting the same limits, and having numerator and denominator degrees at most 2. This leaves (Nåsell 1978, Appendix) the lower bounds $L_{\nu,1,0} < L_{\nu,2,0}$ and the upper bounds $U_{\nu,1,1}$ and $U_{\nu,3,0} < U_{\nu,2,0}$, where the inequalities again follow from Theorems 2 and 3 in the reference. Neither of $U_{\nu,1,1}$ and $U_{\nu,3,0}$ dominates the other, as they have different orders of approximation at 0 and ∞. In fact, writing

$$U_{\nu,1,1}(t) = t \frac{P_{\nu,1,1}(t)}{Q_{\nu,1,1}(t)} = t \frac{2(\nu + 2) + t}{4(\nu + 1)(\nu + 2) + 2(\nu + 1)t + t^2}$$

and

$$U_{\nu,3,0}(t) = t \frac{P_{\nu,3,0}(t)}{Q_{\nu,3,0}(t)} = t \frac{\frac{1}{2}(\nu + 1/2) + t}{(\nu + 1/2)(\nu + 1) + \frac{3}{2}(\nu + 1/2)t + t^2},$$

it is readily verified that

$$U_{\nu,1,1}(t) - U_{\nu,3,0}(t) = \frac{(\nu + 5/2)t^2(t - (\nu + 2))}{Q_{\nu,1,1}(t)Q_{\nu,3,0}(t)}.$$

This implies that with $t_\nu = \nu + 2$, $U_{\nu,1,1}(t) < U_{\nu,3,0}(t)$ for $0 < t < t_\nu$, and $U_{\nu,1,1}(t) > U_{\nu,3,0}(t)$ for $t > t_\nu$.

The best lower bound $L_{\nu,2,0}$ which only involves solving a quadratic equation is given by

$$L_{\nu,2,0} = t \frac{(\nu + 3/2) + t}{2(\nu + 1)(\nu + 3/2) + 2(\nu + 1)t + t^2}.$$

Theorem 5 *Let $\nu \geq 0$. Then for all $t > 0$,*

$$L_{\nu,2,0}(t) < G_{\nu+1/2,\nu+3/2}(t)$$

and

$$\min\left(G_{\nu,\nu+2}(t), G_{\nu+1/2,\beta_{SS}(\nu)}(t)\right) < \min\left(U_{\nu,1,1}(t), U_{\nu,3,0}(t)\right).$$

Proof All Nåsell bounds under consideration are of the form

$$t \frac{P(t)}{Q(t)}, \qquad P(t) = t + \gamma, \quad Q(t) = t^2 + \delta_1 t + \delta_0$$

with non-negative coefficients γ, δ_1 and δ_0. All Amos-type bounds $G_{\alpha,\beta}$ under consideration have $\alpha \leq \beta$.

To show that the Amos-type bounds dominate these Nåsell bounds we have to investigate when

$$\frac{P(t)}{Q(t)} - \frac{1}{\alpha + \sqrt{t^2 + \beta^2}}$$

has no zeros on $(0, \infty)$, or equivalently, when

$$\Delta(t) = (Q(t) - \alpha P(t))^2 - \left(t^2 + \beta^2\right) P(t)^2$$

has no zeros on $(0, \infty)$. Note that

$$\Delta(t) = \left(Q(t) - \left(\alpha + \sqrt{t^2 + \beta^2}\right) P(t)\right) \left(Q(t) - \left(\alpha - \sqrt{t^2 + \beta^2}\right) P(t)\right),$$

where the second term is always positive for $t > 0$ and $\alpha \le \beta$. Hence, $\Delta(t) > 0$ for all $t > 0$ is equivalent to $Q(t) - (\alpha + \sqrt{t^2 + \beta^2}) P(t) > 0$, or $P(t)/Q(t) < 1/(\alpha + \sqrt{t^2 + \beta^2})$ for all $t > 0$; $\Delta(t) < 0$ for all $t > 0$ is equivalent to $P(t)/Q(t) > 1/(\alpha + \sqrt{t^2 + \beta^2})$ for all $t > 0$.

Writing

$$Q(t) - \alpha P(t) = t^2 + (\delta_1 - \alpha)t + (\delta_0 - \alpha\gamma) = t^2 + \omega_1 t + \omega_0,$$

we obtain

$$\begin{aligned}
\Delta(t) &= (t^2 + \omega_1 t + \omega_0)^2 - (t^2 + \beta^2)(t + \gamma)^2 \\
&= 2(\omega_1 - \gamma)t^3 + (\omega_1^2 + 2\omega_0 - \beta^2 - \gamma^2)t^2 + 2(\omega_1\omega_0 - \beta^2\gamma)t + (\omega_0^2 - \beta^2\gamma^2).
\end{aligned}$$

Comparing $L_{\nu,2,0}$ to $G_{\nu+1/2,\nu+3/2}$, we have $\gamma = \nu+3/2$, $\delta_0 = 2(\nu+1)(\nu+3/2)$ and $\delta_1 = 2(\nu+1)$, from which $\omega_0 = \delta_0 - \alpha\gamma = 2(\nu+1)(\nu+3/2) - (\nu+1/2)(\nu+3/2) = (\nu + 3/2)^2 = \beta^2$ and $\omega_1 = \delta_1 - \alpha = \nu + 3/2 = \beta$, giving

$$\begin{aligned}
2(\omega_1 - \gamma) &= 0, \\
\omega_1^2 + 2\omega_0 - \beta^2 - \gamma^2 &= \beta^2, \\
2(\omega_1\omega_0 - \beta^2\gamma) &= 0, \\
\omega_0^2 - \beta^2\gamma^2 &= 0
\end{aligned}$$

so that $\Delta(t) = \beta^2 t^2 > 0$ for all $t > 0$. Hence, we have $L_{\nu,1,0}(t) < L_{\nu,2,0}(t) < G_{\nu+1/2,\nu+3/2}(t)$ for all $t > 0$.

Comparing $U_{\nu,1,1}$ to $G_{\nu,\nu+2}$, we have $\gamma = 2(\nu + 2) = 2\beta$, $\delta_0 = 4(\nu + 1)(\nu + 2)$ and $\delta_1 = 2(\nu + 1)$, from which $\omega_0 = \delta_0 - \alpha\gamma = 4(\nu + 1)(\nu + 2) - 2\nu(\nu + 2) = 2(\nu + 2)^2 = 2\beta^2$ and $\omega_1 = \delta_1 - \alpha = 2(\nu + 1) - \nu = \nu + 2 = \beta$, giving

$$2(\omega_1 - \gamma) = -2\beta,$$
$$\omega_1^2 + 2\omega_0 - \beta^2 - \gamma^2 = 0,$$
$$2(\omega_1\omega_0 - \beta^2\gamma) = 0,$$
$$\omega_0^2 - \beta^2\gamma^2 = 0,$$

so that $\Delta(t) = -2\beta t^3 < 0$ for all $t > 0$. Hence, $G_{\nu,\nu+2}(t) < U_{\nu,1,1}(t)$ for all $t > 0$.

Finally, comparing $U_{\nu,3,0}$ to $G_{\nu+1/2,\beta_{SS}(\nu)}$, we have $\gamma = \frac{1}{2}(\nu + 1/2) = \alpha/2$, $\delta_0 = (\nu + 1/2)(\nu + 1) = \alpha(\alpha + 1/2)$. $\delta_1 = \frac{3}{2}(\nu + 1/2) = 3\alpha/2$, from which $\omega_0 = \delta_0 - \alpha\gamma = \alpha(\alpha + 1/2) - \alpha^2/2 = \alpha(\alpha + 1)/2 = \beta^2/2$ and $\omega_1 = \delta_1 - \alpha = 3\alpha/2 - \alpha = \alpha/2$, giving

$$2(\omega_1 - \gamma) = 0,$$
$$\omega_1^2 + 2\omega_0 - \beta^2 - \gamma^2 = 0,$$
$$2(\omega_1\omega_0 - \beta^2\gamma) = -\alpha\beta^2/2,$$
$$\omega_0^2 - \beta^2\gamma^2 = \beta^2(\beta^2 - \alpha^2)/4,$$

so that with $\beta^2 - \alpha^2 = (\nu + 1/2)(\nu + 3/2) - (\nu + 1/2)^2 = \nu + 1/2 = \alpha$ we get

$$\Delta(t) = -\frac{\alpha\beta^2}{2}t + \frac{\beta^2\alpha}{4} = \frac{\beta^2\alpha}{4}(1 - 2t),$$

which is negative for $t > 1/2$, so that $G_{\nu+1/2,\beta_{SS}(\nu)}(t) < U_{\nu,3,0}(t)$ for all $t > 1/2$. As $U_{\nu,1,1}(t) < U_{\nu,3,0}(t)$ for $0 < t < t_\nu = \nu + 2$, we infer that

$$\min\left(G_{\nu,\nu+2}(t), G_{\nu+1/2,\beta_{SS}(\nu)}(t)\right) < \min\left(U_{\nu,1,1}(t), U_{\nu,3,0}(t)\right)$$

for all $t > 0$, and the proof is complete.

As the Nåsell bounds considered in the previous theorem are dominated by the Amos-type bounds used for R_ν, the same must be true for the respective inverses.

Corollary 3 *Let $\nu \geq 0$ and $0 \leq \rho < 1$. Then*

$$R_\nu^{-1}(\rho) \leq G_{\nu+1/2,\nu+3/2}^{-1}(\rho) < L_{\nu,2,0}^{-1}(\rho)$$

and

$$\max\left(U_{\nu,1,1}^{-1}(\rho), U_{\nu,3,0}^{-1}(\rho)\right) \leq \max\left(G_{\nu,\nu+2}^{-1}(\rho), G_{\nu+1/2,\beta_{SS}(\nu)}^{-1}(\rho)\right) \leq R_\nu^{-1}(\rho).$$

Proof Straightforward from the previous theorem and Theorem 1.

Table 1 Number of iterations required by `uniroot()` (top) and by the Newton-Fourier algorithm (bottom) using the Tanabe et al. (2007) bounds (T07) and the newly established bounds (New) for initialization

T07	New					
	0	1	2	3	4	5
1	4	3	0	0	0	0
2	2	22	32	2	0	0
3	0	5	46	36	3	0
4	0	0	5	14	9	0
5	0	0	4	3	1	1
T07	New					
	1	2				
1	87	0				
2	52	45				
3	0	8				

4 Numerical comparisons

In the following we compare the number of iterations required to reach convergence by two different algorithms based on nested intervals for finding roots using (1) the Tanabe et al. (2007) and (2) the newly established bounds for initialization. The two algorithms used are (a) a one-dimensional root finding algorithm as implemented in function `uniroot()` available in R (R Core Team 2013) and (b) the variant of the Newton-Fourier method for the case of strictly increasing concave functions (see e.g., Atkinson 1989, pp. 62–64). Concavity of R_ν can be established by using that R_ν is the pointwise minimum of a set of Amos-type functions (see Hornik and Grün 2013, Theorem 11). It is straightforward to show that the second derivative of these Amos-type functions is non-positive and hence R_ν is concave, because it is the pointwise minimum of concave functions.

The numbers of iterations required are determined for a range of different d and κ values which are selected similar to those used in Sra (2012), i.e., $\kappa = 100, 500, 1,000, 5,000, 10,000, 20,000, 50,000, 10,000$ and $d = 500, 1,000, 5,000, 10,000, 20,000, 10,000$. Each value of d is combined with each value of κ and they are used to determine the corresponding ρ. For the given d and implied ρ the root finding algorithms are employed to determine κ again. For `uniroot()` convergence is assessed using argument `tol` and the Newton-Fourier algorithm is stopped if the relative distance of the midpoint to the endpoints of the interval is smaller than `tol`. This pre-specified precision value `tol` is varied and set to values 10^e for exponents $e \in \{-6, -8, -10, -12\}$.

The results are given in Table 1. For the Newton-Fourier algorithm the newly established bounds always require the same or a smaller number of iterations. In 31 % of the cases one iteration less is required, reducing the number of iterations required by 60 or 20 %. When using `uniroot()`, there is an improvement achievement in 55 % of the cases, the same number of iterations are required in 42 % of the cases and only in 3 % of the cases a deterioration occurs and one iteration more is required when using the newly established bounds for initialization.

Overall it has to be noted that the number of iterations is rather small to reach convergence when either of the two sets of intervals is used for initialization. However, the general tendency to require even less iterations of the newly established bounds will nevertheless be of interest from a computational point of view if these roots are solved repeatedly, which is required when for example the expectation-maximization algorithm is used to estimate mixtures of von Mises–Fisher distributions (Banerjee et al. 2005).

References

Amos DE (1974) Computation of modified Bessel functions and their ratios. Math Comput 28(125):239–251. http://www.jstor.org/stable/2005830

Atkinson KE (1989) An introduction to numerical analysis, 2nd edn. Wiley, New York

Banerjee A, Dhillon IS, Ghosh J, Sra S (2005) Clustering on the unit hypersphere using von Mises–Fisher distributions. J Mach Learn Res 6:1345–1382. http://jmlr.csail.mit.edu/papers/v6/banerjee05a.html

Dhillon IS, Sra S (2003) Modeling data using directional distributions. Tech. Rep. TR-03-06, Department of Computer Sciences, The University of Texas at Austin. ftp://ftp.cs.utexas.edu/pub/techreports/tr03-06.ps.gz

Gautschi W, Slavik J (1978) On the computation of modified Bessel function ratios. Math Comput 32(143):865–875. http://www.jstor.org/stable/2006491

Hornik K, Grün B Amos-type bounds for modified Bessel function ratios. J Math Anal Appl. doi:10.1016/j.jmaa.2013.05.070

Mardia KV, Jupp PE (1999) Directional statistics. Probability and Statistics. Chichester, UK

Nåsell I (1978) Rational bounds for ratios of modified Bessel functions. SIAM J Math Anal 9(1):1–11. doi:10.1137/0509001

R Core Team (2013) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. http://www.R-project.org/

Schou G (1978) Estimation of the concentration parameter in von Mises-Fisher distributions. Biometrika 65(1):369–377

Simpson HC, Spector SJ (1984) Some monotonicity results for ratios of modified Bessel functions. Quart Appl Math 42(1):95–98

Song H, Liu J, Wang G (2012) High-order parameter approximation for von Mises-Fisher distributions. Appl Math Comput 218(24):11880–11890. doi:10.1016/j.amc.2012.05.050

Sra S (2012) A short note on parameter approximation for von Mises-Fisher distributions: and a fast implementation of $I_s(x)$. Comput Stat 27(1):177–190. doi:10.1007/s00180-011-0232-x

Tanabe A, Fukumizu K, Oba S, Takenouchi T, Ishii S (2007) Parameter estimation for von Mises-Fisher distributions. Comput Stat 22:145–157. doi:10.1007/s00180-007-0030-7

Tretter MJ, Walster GW (1980) Further comments on the computation of modified Bessel function ratios. Math Comput 35(151):937–939. http://www.jstor.org/stable/2006205

Watson GN (1995) A treatise on the theory of bessel functions, 2nd edn. Cambridge mathematical library, Cambridge University Press. doi:10.2277/0521483913; http://www.cambridge.org/9780521483919

Yuan L, Kalbfleisch JD (2000) On the Bessel distribution and related problems. Ann Inst Stat Math 52(3):438–447