Volumes of unit balls of mixed sequence spaces

Henning Kempka,[*] Jan Vybíral[†]

July 6, 2015

Abstract

The volume of the unit ball of the Lebesgue sequence space ℓ_p^m is very well known since the times of Dirichlet. We calculate the volume of the unit ball in the mixed norm $\ell_q^n(\ell_p^m)$, whose special cases are nowadays popular in machine learning under the name of group lasso. We consider the real as well as the complex case. The result is given by a closed formula involving the gamma function, only slightly more complicated than the one of Dirichlet. We close by an overview of open problems.

1 Introduction

Exact formulas and estimates of volumes of convex bodies play a central role in functional analysis and geometry of Banach spaces with numerous connections to probability theory, approximation theory and other areas of mathematics. We refer to the classical book [17] for an extensive treatment.

Unfortunately, for most of the bodies it is out of reach to express their volume by simple formulas. The most important exception from this rule is the volume of the unit ball of the Lebesgue sequence space

$$B_p^m(\mathbb{R}) = \{x = (x_1, \ldots, x_m) \in \mathbb{R}^m : \sum_{j=1}^m |x_j|^p \le 1\}$$

for $0 < p < \infty$ (with $B_\infty^m(\mathbb{R})$ being just the cube $[-1,1]^m$ with volume 2^m). The volume of $B_p^m(\mathbb{R})$ is very well known since the time of Dirichlet, cf. [9, 13]. It can be expressed directly with the help of the gamma function as

$$\text{vol}(B_p^m(\mathbb{R})) = 2^m \frac{\Gamma\left(\frac{1}{p} + 1\right)^m}{\Gamma\left(\frac{m}{p} + 1\right)}, \tag{1}$$

1

where $1/p$ is interpreted as zero for $p = \infty$.

The approach of Dirichlet is based on inductive formula for evaluation of multivariate integrals. Alternatively, a noninductive way to calculate $\mathrm{vol}(B_p^m(\mathbb{R}))$ is based on Gaussian integrals. We review briefly this approach using the notation of [19]. For $0 < p < \infty$ we define

$$\Delta_p^m = \{x \in \mathbb{R}^m : x_j \geq 0 \text{ for all } j = 1, \ldots, m \text{ and } \|x\|_p = 1\}$$

and the normalized cone measure

$$\mu_p(A) = \frac{\lambda([0,1] \cdot A)}{\lambda([0,1] \cdot \Delta_p^m)}, \qquad A \subset \Delta_p^m,$$

where λ is the usual Lebesgue measure in $\mathbb{R}^m$. Applying the *polar decomposition identity* [2]

$$\frac{\int_{\mathbb{R}_+^m} f(x) d\lambda(x)}{\lambda([0,1] \cdot \Delta_p^m)} = m \int_0^\infty r^{m-1} \int_{\Delta_p^m} f(rx) d\mu_p(x) dr$$

to $f(x) = e^{-x_1^p - \cdots - x_m^p}$, one gets

$$\frac{\left(\int_0^\infty e^{-t^p} dt\right)^m}{2^{-m} \mathrm{vol}(B_p^m(\mathbb{R}))} = m \int_0^\infty r^{m-1} e^{-r^p} dr.$$

After the substitution $s = t^p$ and employing the definition of the gamma function, the formula (1) follows.

The results on volumes of unit balls of finite dimensional spaces can be usually translated into the language of entropy numbers of their embeddings. Furthermore, by the results of Carl and Triebel [3, 4], they have implications on the decay of different approximation widths as well as on the decay of eigenvalues of compact operators. Finally, by the use of discretization techniques, these results can be exploited to study these quantities also for embeddings of function spaces, see [23] for an overview of this approach and many references. When applying this approach to weighted function spaces or function spaces with dominating mixed smoothness, entropy numbers of embeddings of mixed Lebesgue spaces $\ell_q^n(\ell_p^m)$ start to play a role, cf. [21, 22]. We refer also to [7, 12, 14, 15] for other results on finite dimensional sequence spaces with mixed norm.

Since the advent of Lasso [20] in 1996 and its success in analysis of high dimensional data, there was a new wave of interest in geometry of finite dimensional sequence spaces. Ten years later it was exhibited in the theory of compressed sensing [5, 6, 10] that the minimization of the ℓ_1-norm can be used for recovery of sparse data from a small amount of random linear measurements. To recover data with more complex sparsity patterns, other ℓ_1-based norms were also studied. The most important functional in this connection is given by

$$\min_{x \in \mathbb{R}^{n \times m}} \sum_{j=1}^n \left(\sum_{k=1}^m |x_{j,k}|^2\right)^{1/2},$$

2

which will be denoted by $\|x\|_{2,1}$ later on. Minimization of this norm is usually denoted by *group lasso* [1, 16] and is used to recover data with *block sparsity*, i.e. double sequences $(x_{j,k})$, for which the set

$$\{j \in \{1, \ldots, n\} : \exists\, k \in \{1, \ldots, m\} \text{ such that } x_{j,k} \neq 0\}$$

has cardinality much smaller than n, cf. [24]. The rest of the paper is structured as follows. In Section 2 we evaluate an auxiliary multivariate integral, in Section 3 we prove our main result, Theorem 3, extend the result to the complex case and discuss few special cases in detail. We close in Section 4 with a list of open problems.

2 Auxiliary integrals

In this section we want to calculate the value of the integrals

$$A_m(\alpha, \beta) = \int\limits_{\substack{a \geq 0 \\ 0 \leq a_1 + \cdots + a_m \leq 1}} (1 - a_1 - \cdots - a_m)^\beta a_1^\alpha \ldots a_m^\alpha \, da$$

for $m \in \mathbb{N}$ and $\alpha, \beta > -1$. The integration in this notation is with respect to the set

$$\{a = (a_1, \ldots, a_m) \in \mathbb{R}^m : a_j \geq 0 \text{ for all } j = 1, \ldots, m \text{ and } 0 \leq a_1 + \cdots + a_m \leq 1\}.$$

For that reason we use for $t > 0$ the gamma function

$$\Gamma(t) = \int\limits_0^\infty x^{t-1} e^{-x} dx$$

with the well known properties $\Gamma(t + 1) = t\Gamma(t)$, $\Gamma(1) = 1$ and $\Gamma(1/2) = \sqrt{\pi}$. Furthermore, we need the beta function and the following identity

$$B(\alpha, \beta) = \int\limits_0^1 t^{\alpha-1}(1 - t)^{\beta-1} dt = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)},$$

which holds for all $\alpha, \beta > 0$, see [8]. Now, we are ready to prove the following identity for $A_m(\alpha, \beta)$.

Theorem 1. *Let $m \in \mathbb{N}$ and $\alpha, \beta > -1$ then*

$$A_m(\alpha, \beta) = \frac{\Gamma(\alpha + 1)^m \Gamma(\beta + 1)}{\Gamma(m(\alpha + 1) + \beta + 1)}.$$

By choosing $\alpha = \beta > -1$ we especially have

$$A_m(\alpha, \alpha) = \frac{\Gamma(\alpha + 1)^{m+1}}{\Gamma((m + 1)(\alpha + 1))}.$$

3

Proof. First we can observe

$$A_1(\alpha, \beta) = \int\limits_0^1 (1-a)^\beta a^\alpha da = B(\alpha+1, \beta+1) = \frac{\Gamma(\alpha+1)\Gamma(\beta+1)}{\Gamma(\alpha+\beta+2)}. \qquad (2)$$

The rest of the proof follows from a recursion formula. We use the substitution $(b_1, b_2, \ldots, b_{m-1}, b_m) = (a_1, a_2, \ldots, a_{m-1}, a_1 + \cdots + a_m)$ with Jacobian equal to one. It follows

$$A_m(\alpha, \beta) = \int\limits_{\substack{a \geq 0 \\ 0 \leq a_1 + \cdots + a_m \leq 1}} (1 - a_1 - \cdots - a_m)^\beta a_1^\alpha \ldots a_m^\alpha da$$

$$= \int\limits_{\substack{b \geq 0 \\ 0 \leq b_1 + \cdots + b_{m-1} \leq b_m \leq 1}} (1 - b_m)^\beta b_1^\alpha \ldots b_{m-1}^\alpha (b_m - b_1 - \cdots - b_{m-1})^\alpha db$$

$$= \int\limits_0^1 (1 - b_m)^\beta \int\limits_{\substack{b' \geq 0 \\ 0 \leq b_1 + \cdots + b_{m-1} \leq b_m}} b_1^\alpha \ldots b_{m-1}^\alpha (b_m - b_1 - \cdots - b_{m-1})^\alpha db' db_m,$$

where $b' = (b_1, \ldots, b_{m-1})$. Then we substitute $c_1 = b_1/b_m, \ldots, c_{m-1} = b_{m-1}/b_m$ with $\left| \frac{db'}{dc} \right| = b_m^{m-1}$, which gives

$$= \int\limits_0^1 (1 - b_m)^\beta b_m^{m\alpha+m-1} db_m \int\limits_{\substack{c \geq 0 \\ 0 \leq c_1 + \cdots + c_{m-1} \leq 1}} c_1^\alpha \ldots c_{m-1}^\alpha (1 - c_1 - \cdots - c_{m-1})^\alpha dc$$

$$= B(m(\alpha+1), \beta+1) A_{m-1}(\alpha, \alpha).$$

So we have the identity

$$A_m(\alpha, \beta) = A_{m-1}(\alpha, \alpha) \frac{\Gamma(m(\alpha+1))\Gamma(\beta+1)}{\Gamma(m(\alpha+1) + \beta + 1)}. \qquad (3)$$

The rest of the proof follows by induction from (2) and (3). $\qquad \square$

3   The volume of the unit ball in $\ell_q^n(\ell_p^m)$

We are interested in the volume of the unit ball of the mixed sequence space $\ell_q^n(\ell_p^m)$. First of all we treat the real case.

4

3.1 Unit ball in $\ell_q^n(\ell_p^m(\mathbb{R}))$

If $x = (x_{j,k})_{j,k=1}^{n,m}$ are real numbers, then we define for all $0 < p, q \leq \infty$

$$\|x\|_{p,q} = \left(\sum_{j=1}^{n} \left(\sum_{k=1}^{m} |x_{j,k}|^p \right)^{q/p} \right)^{1/q} \quad \text{and} \quad B_{p,q}^{m,n}(\mathbb{R}) = \{x \in \mathbb{R}^{n \times m} : \|x\|_{p,q} \leq 1\}$$

with appropriate modifications if p or q is equal to infinity.
In the non-mixed case of $\ell_p^m(\mathbb{R})$ the following result is well known [17, (1.17)].

Proposition 2. *Let $m \in \mathbb{N}$ and $0 < p \leq \infty$ then*

$$\text{vol}(B_p^m(\mathbb{R})) = 2^m \frac{\Gamma\left(\frac{1}{p}+1\right)^m}{\Gamma\left(\frac{m}{p}+1\right)}. \tag{4}$$

For the mixed spaces $\ell_q^n(\ell_p^m(\mathbb{R}))$ the following theorem is our main result.

Theorem 3. *Let $m, n \in \mathbb{N}$ and $0 < p, q \leq \infty$ then*

$$\text{vol}(B_{p,q}^{m,n}(\mathbb{R})) = 2^{mn} \frac{\Gamma\left(\frac{1}{p}+1\right)^{mn} \Gamma\left(\frac{m}{q}+1\right)^n}{\Gamma\left(\frac{nm}{q}+1\right) \Gamma\left(\frac{m}{p}+1\right)^n}. \tag{5}$$

Proof. We give the proof for $\max(p,q) < \infty$. The modifications necessary in case of $p = \infty$ or $q = \infty$ are described in the subsection on special cases below. First of all we restrict the volume to positive $x \geq 0$ (meaning $x_{j,k} \geq 0$ for all j,k)

$$\text{vol}(B_{p,q}^{m,n}(\mathbb{R})) = \int_{B_{p,q}^{m,n}(\mathbb{R})} 1dx = 2^{mn} \int_{B_{p,q}^{m,n}(\mathbb{R}), x \geq 0} 1dx$$

and using the substitution $s_{j,k} = x_{j,k}^p$ we obtain

$$= \frac{2^{mn}}{p^{mn}} \int_{s \in A_{p,q}^{m,n}(1)} \prod_{j=1}^{n} \prod_{k=1}^{m} s_{j,k}^{1/p-1} ds, \tag{6}$$

where the integral is over $A_{p,q}^{m,n}(t) = \left\{ s \in \mathbb{R}^{n \times m} : s \geq 0 \text{ and } \sum_{j=1}^{n} \left(\sum_{k=1}^{m} s_{j,k} \right)^{q/p} \leq t \right\}$.

Now, we can rewrite

$$\text{vol}(B_{p,q}^{m,n}(\mathbb{R})) = \frac{2^{mn}}{p^{mn}} \int_{\substack{s \geq 0 \\ 0 \leq s_{n,1} + \cdots + s_{n,m} \leq 1}} (s_{n,1} \ldots s_{n,m})^{1/p-1} \int_{s' \in A_{p,q}^{m,n-1}(1-(s_{n,1}+\cdots+s_{n,m})^{q/p})} \prod_{j=1}^{n-1} \prod_{k=1}^{m} s_{j,k}^{1/p-1} ds' \, ds_n,$$

where $s' = (s_{j,k} : j = 1, \ldots, n-1$ and $k = 1, \ldots, m)$ and $s_n = (s_{n,1}, \ldots, s_{n,m})$. We substitute for $j < n$ with $\sigma_{j,k} = \frac{s_{j,k}}{[1-(s_{n,1}+\cdots+s_{n,m})^{q/p}]^{p/q}}$ and get

$$= \frac{2^{mn}}{p^{mn}} \int\limits_{\substack{s \geq 0 \\ 0 \leq s_{n,1}+\cdots+s_{n,m} \leq 1}} (s_{n,1} \ldots s_{n,m})^{1/p-1} \left[1 - (s_{n,1} + \cdots + s_{n,m})^{q/p}\right]^{\frac{m(n-1)}{q}} ds_n \times$$

$$\times \int\limits_{\sigma \in A_{p,q}^{m,n-1}(1)} \prod_{j=1}^{n-1} \prod_{k=1}^{m} \sigma_{j,k}^{1/p-1} d\sigma$$

$$= \text{vol}(B_{p,q}^{m,n-1}(\mathbb{R})) \frac{2^m}{p^m} \int\limits_{\substack{s \geq 0 \\ 0 \leq s_1+\cdots+s_m \leq 1}} (s_1 \ldots s_m)^{1/p-1} \left[1 - (s_1 + \cdots + s_m)^{q/p}\right]^{\frac{m(n-1)}{q}} ds,$$

(7)

where the last equality follows from comparison to (6). Now, we evaluate the last integral in (7) putting $\lambda_1 = s_1, \ldots, \lambda_{m-1} = s_{m-1}$ and $\lambda_m = s_1 + \cdots + s_m$

$$\int\limits_{\substack{s \geq 0 \\ 0 \leq s_1+\cdots+s_m \leq 1}} (s_1 \ldots s_m)^{1/p-1} \left[1 - (s_1 + \cdots + s_m)^{q/p}\right]^{\frac{m(n-1)}{q}} ds$$

$$= \int\limits_{\substack{\lambda \geq 0 \\ 0 \leq \lambda_1+\cdots+\lambda_{m-1} \leq \lambda_m \leq 1}} (\lambda_1 \ldots \lambda_{m-1})^{1/p-1} (\lambda_m - (\lambda_1 + \cdots \lambda_{m-1}))^{1/p-1} (1 - \lambda_m^{q/p})^{\frac{m(n-1)}{q}} d\lambda$$

$$= \int_0^1 (1 - \lambda_m^{q/p})^{\frac{m(n-1)}{q}} \int\limits_{\substack{\lambda' \geq 0 \\ 0 \leq \lambda_1+\cdots+\lambda_{m-1} \leq \lambda_m}} (\lambda_1 \ldots \lambda_{m-1})^{1/p-1} (\lambda_m - (\lambda_1 + \cdots \lambda_{m-1}))^{1/p-1} d\lambda' \, d\lambda_m$$

and substituting $\mu_j = \frac{\lambda_j}{\lambda_m}$ for $j \leq m-1$ we obtain

$$= \int_0^1 (1 - \lambda_m^{q/p})^{\frac{m(n-1)}{q}} \lambda_m^{m/p-1} d\lambda_m \int\limits_{\substack{\mu \geq 0 \\ 0 \leq \mu_1+\cdots+\mu_{m-1} \leq 1}} (\mu_1 \ldots \mu_{m-1})^{1/p-1} (1 - (\mu_1 + \cdots \mu_{m-1}))^{1/p-1} d\mu$$

$$= A_{m-1}(1/p-1, 1/p-1) \int_0^1 (1 - \lambda_m^{q/p})^{\frac{m(n-1)}{q}} \lambda_m^{m/p-1} d\lambda_m. \tag{8}$$

Finally, we use $t = \lambda_m^{q/p}$ and see that the last integral equals

$$\int\limits_0^1 (1 - \lambda_m^{q/p})^{\frac{m(n-1)}{q}} \lambda_m^{m/p-1} d\lambda_m = \frac{p}{q} \int\limits_0^1 (1-t)^{\frac{m(n-1)}{q}} t^{\frac{m}{q}-1} dt$$

$$= \frac{p}{q} B\left(\frac{m}{q}, \frac{m(n-1)}{q} + 1\right) = \frac{p}{q} \frac{\Gamma(\frac{m}{q})\Gamma(\frac{m(n-1)}{q} + 1)}{\Gamma(\frac{mn}{q} + 1)} \quad (9)$$

Using the recursive formula (7) with (8) and (9) and Theorem 1 we obtain

$$\mathrm{vol}(B_{p,q}^{m,n}(\mathbb{R})) = \mathrm{vol}(B_{p,q}^{m,n-1}(\mathbb{R})) \frac{2^m}{p^m} A_{m-1}\left(\frac{1}{p} - 1, \frac{1}{p} - 1\right) \frac{p}{q} \frac{\Gamma(\frac{m}{q})\Gamma(\frac{m(n-1)}{q} + 1)}{\Gamma(\frac{mn}{q} + 1)}$$

$$= \mathrm{vol}(B_{p,q}^{m,n-1}(\mathbb{R})) \frac{2^m}{p^m} \frac{p}{q} \frac{\Gamma(\frac{1}{p})^m}{\Gamma(\frac{m}{p})} \frac{\Gamma(\frac{m}{q})\Gamma(\frac{m(n-1)}{q} + 1)}{\Gamma(\frac{mn}{q} + 1)}$$

$$= \mathrm{vol}(B_{p,q}^{m,n-1}(\mathbb{R})) 2^m \frac{mp}{mq} \frac{\Gamma(\frac{m}{q})}{\Gamma(\frac{m}{p})} \Gamma\left(\frac{1}{p} + 1\right)^m \frac{\Gamma(\frac{m(n-1)}{q} + 1)}{\Gamma(\frac{mn}{q} + 1)}$$

and using (7) $n - 2$ times we have

$$= \mathrm{vol}(B_{p,q}^{m,1}(\mathbb{R})) \left[2^m \frac{\Gamma(\frac{m}{q} + 1)}{\Gamma(\frac{m}{p} + 1)} \Gamma\left(\frac{1}{p} + 1\right)^m\right]^{n-1} \frac{\Gamma(\frac{m}{q} + 1)}{\Gamma(\frac{mn}{q} + 1)}.$$

$$(10)$$

We observe that $B_{p,q}^{m,1}(\mathbb{R}) = B_p^m(\mathbb{R})$ and using Proposition 2 we finally get

$$\mathrm{vol}(B_{p,q}^{m,n}(\mathbb{R})) = 2^{m(n-1)} 2^m \frac{\Gamma(\frac{1}{p} + 1)^m}{\Gamma(\frac{m}{p} + 1)} \frac{\Gamma(\frac{m}{q} + 1)^{n-1}}{\Gamma(\frac{m}{p} + 1)^{n-1}} \Gamma\left(\frac{1}{p} + 1\right)^{m(n-1)} \frac{\Gamma(\frac{m}{q} + 1)}{\Gamma(\frac{mn}{q} + 1)}$$

$$= 2^{mn} \frac{\Gamma(\frac{m}{q} + 1)^n}{\Gamma(\frac{m}{p} + 1)^n} \frac{\Gamma(\frac{1}{p} + 1)^{mn}}{\Gamma(\frac{mn}{q} + 1)},$$

which finishes the proof. $\qquad\square$

3.2 Alternative proof

We present an alternative proof of (5), which was communicated to us by Prof. Franck Barthe shortly after the first version of this paper was finished.

Let $K \subset \mathbb{R}^d$ be a symmetric convex body with non-empty interior and let $\|\cdot\|_K$ be the corresponding norm (i.e. its Minkowski functional).

Modifying the classical computation from convex geometry (cf. [17, page 11]) we get for smooth non-negative function f on $[0, \infty)$ with fast decay at

infinity

$$\int_{\mathbb{R}^d} f(\|x\|_K)dx = -\int_{\mathbb{R}^d} \int_{\|x\|_K}^{\infty} f'(t)dt\,dx = -\int_0^{\infty} \int_{x:\|x\|_K \leq t} 1\,dx\,f'(t)dt$$

$$= -\int_0^{\infty} \operatorname{vol}(tK)f'(t)dt = -\operatorname{vol}(K)\int_0^{\infty} t^d f'(t)dt \qquad (11)$$

$$= \operatorname{vol}(K) \cdot \int_0^{\infty} dt^{d-1}f(t)dt.$$

Plugging $f(t) = e^{-t^p}$ into (11), we obtain in particular

$$\int_{\mathbb{R}^d} e^{-\|x\|_K^p}dx = \operatorname{vol}(K) \cdot \int_0^{\infty} dt^{d-1}e^{-t^p}\,dt = \operatorname{vol}(K) \cdot \frac{d}{p} \cdot \int_0^{\infty} s^{d/p-1}e^{-s}ds$$

$$= \frac{d\operatorname{vol}(K)\Gamma(d/p)}{p} = \operatorname{vol}(K)\Gamma(1+d/p). \qquad (12)$$

This formula can already be used to give the volume of the ℓ_p^d-unit ball. Indeed, we get for $K = B_p^d$

$$\operatorname{vol}(B_p^d)\Gamma(1+d/p) = \int_{\mathbb{R}^d} e^{-\|x\|_p^p}dx = \Big(2\int_0^{\infty} e^{-t^p}\,dt\Big)^d = 2^d\Big(\frac{1}{p}\int_0^{\infty} s^{1/p-1}e^{-s}ds\Big)^d$$

$$= 2^d\Big(\frac{\Gamma(1/p)}{p}\Big)^d = 2^d\Gamma(1+1/p)^d, \qquad (13)$$

giving (1) again. Choosing K to be the unit ball of $\ell_q^n(\ell_p^m(\mathbb{R}))$, we get by (12)

$$\Gamma\Big(1+\frac{mn}{q}\Big)\operatorname{vol}(B_{p,q}^{m,n}(\mathbb{R})) = \int_{\mathbb{R}^{n\times m}} e^{-\|x\|_{p,q}^q}dx$$

$$= \int_{\mathbb{R}^{n\times m}} \exp\Big(-\sum_{j=1}^{n}\big(\sum_{k=1}^{m}|x_{j,k}|^p\big)^{q/p}\Big)dx$$

$$= \prod_{j=1}^{n}\int_{\mathbb{R}^m} \exp\Big(-\big(\sum_{k=1}^{m}|x_{j,k}|^p\big)^{q/p}\Big)dx_{j,\cdot} \qquad (14)$$

$$= \Big(\int_{\mathbb{R}^m} \exp\Big(-\big(\sum_{k=1}^{m}|x_k|^p\big)^{q/p}\Big)dx\Big)^n$$

$$= \Big(\int_{\mathbb{R}^m} e^{-\|x\|_p^q}dx\Big)^n = \Big[\Gamma\Big(1+\frac{m}{q}\Big)\operatorname{vol}(B_p^m)\Big]^n,$$

8

where we have used Fubini's theorem and (12) again in the last line. Plugging (13) with m instead of d gives finally

$$\mathrm{vol}(B_{p,q}^{m,n}(\mathbb{R})) = \frac{1}{\Gamma\left(1+\frac{mn}{q}\right)}\left[\Gamma\left(1+\frac{m}{q}\right)\frac{2^m\Gamma(1+1/p)^m}{\Gamma(1+m/p)}\right]^n$$

$$= 2^{mn}\frac{\Gamma\left(\frac{1}{p}+1\right)^{mn}\Gamma\left(\frac{m}{q}+1\right)^n}{\Gamma\left(\frac{nm}{q}+1\right)\Gamma\left(\frac{m}{p}+1\right)^n}.$$

As pointed to us by Aicke Hinrichs, this approach can be easily generalized to calculate the volume of the unit ball of the space $\ell_q^n(X_j)$, where X_j are d_j-dimensional (quasi-)Banach spaces and $\|x\|_{\ell_q^n(X_j)} = \left(\sum_{j=1}^n \|x_j\|_{X_j}^q\right)^{1/q}$ giving

$$\mathrm{vol}(B_{\ell_q^n(X_j)}) = \frac{\prod_{j=1}^n \Gamma(1+\frac{d_j}{q})\,\mathrm{vol}(B_{X_j})}{\Gamma\left(1+\frac{\sum_{j=1}^n d_j}{q}\right)}. \tag{15}$$

Finally, we observe that (12) and (14) can be easily adapted to calculate the volume of an anisotropic unit ball

$$B_{p_1,\ldots,p_n} = \{x \in \mathbb{R}^n : \|x\|_{p_1,\ldots,p_n} = |x_1|^{p_1} + \cdots + |x_n|^{p_n} \leq 1\}.$$

What we need is

$$\mathrm{vol}(\{x \in \mathbb{R}^n : \|x\|_{p_1,\ldots,p_n} \leq t\}) = \mathrm{vol}(\{x \in \mathbb{R}^n : |x_1|^{p_1} + \cdots + |x_n|^{p_n} \leq t\})$$

$$= \mathrm{vol}\left(\left\{x \in \mathbb{R}^n : \left|\frac{x_1}{t^{1/p_1}}\right|^{p_1} + \cdots + \left|\frac{x_n}{t^{1/p_n}}\right|^{p_n} \leq 1\right\}\right)$$

$$= \mathrm{vol}(D_\tau B_{p_1,\ldots,p_n}) \tag{16}$$

$$= t^{1/p_1+\cdots+1/p_n}\,\mathrm{vol}(B_{p_1,\ldots,p_n})$$

for every $t > 0$. Here, D_τ denotes a diagonal matrix with $t^{1/p_1},\ldots,t^{1/p_n}$ on the diagonal. With the help of (16), we get the analogue of (11)

$$\int_{\mathbb{R}^d} f(\|x\|_{p_1,\ldots,p_n})dx = -\int_{\mathbb{R}^d}\int_{\|x\|_{p_1,\ldots,p_n}}^\infty f'(t)dtdx = -\int_0^\infty \int_{x:\|x\|_{p_1,\ldots,p_n}\leq t} 1\,dx f'(t)dt$$

$$= -\,\mathrm{vol}(B_{p_1,\ldots,p_n})\int_0^\infty t^{1/p_1+\ldots 1/p_n}f'(t)dt \tag{17}$$

$$= \mathrm{vol}(B_{p_1,\ldots,p_n})\left(\frac{1}{p_1}+\cdots+\frac{1}{p_n}\right)\cdot\int_0^\infty t^{1/p_1+\ldots 1/p_n-1}f(t)dt,$$

where f is a smooth non-negative function on $[0,\infty)$ with fast decay at infinity. Now, using $f(t) = e^{-t}$ gives us

$$\mathrm{vol}(B_{p_1,\ldots,p_n}) = 2^n\frac{\Gamma(1+1/p_1)\cdot\ldots\cdot\Gamma(1+1/p_n)}{\Gamma(1+1/p_1+\cdots+1/p_n)}, \tag{18}$$

which coincides with the result obtained by Dirichlet in [9]. Although one could easily combine for example (18) with (15) to give volumes of more and more complicated bodies, we do not go into details and leave this to the interested reader.

3.3 Unit ball in $\ell_q^n(\ell_p^m(\mathbb{C}))$

If $z = (z_{j,k})_{j,k=1}^{n,m}$ are complex numbers, then we define for all $0 < p, q \le \infty$

$$\|z\|_{p,q} = \left(\sum_{j=1}^n \Big(\sum_{k=1}^m |z_{j,k}|^p \Big)^{q/p} \right)^{1/q} \quad \text{and} \quad B_{p,q}^{m,n}(\mathbb{C}) = \{z \in \mathbb{C}^{n \times m} : \|z\|_{p,q} \le 1\}.$$

Since $\mathbb{C}$ maybe identified with $\mathbb{R}^2$ we can reformulate

$$B_{p,q}^{m,n}(\mathbb{C}) = \left\{ u, v \in \mathbb{R}^{n \times m} : \sum_{j=1}^n \Big(\sum_{k=1}^m (u_{j,k}^2 + v_{j,k}^2)^{p/2} \Big)^{q/p} \le 1 \right\}.$$

In the non-mixed case the following result holds.

Proposition 4 (Proposition 3.2.1 in [11]). *Let $0 < p \le \infty$ then the volume of the unit ball in $\ell_p^m(\mathbb{C})$ equals*

$$\mathrm{vol}(B_p^m(\mathbb{C})) = \pi^m \frac{\Gamma\left(\frac{2}{p} + 1\right)^m}{\Gamma\left(\frac{2m}{p} + 1\right)} = \left(\frac{\pi}{2}\right)^m \mathrm{vol}\left(B_{\frac{p}{2}}^m(\mathbb{R})\right).$$

A similar result holds in the mixed case.

Theorem 5. *Let $m, n \in \mathbb{N}$ and $0 < p, q \le \infty$ then*

$$\mathrm{vol}(B_{p,q}^{m,n}(\mathbb{C})) = \left(\frac{\pi}{2}\right)^{mn} \mathrm{vol}\left(B_{\frac{p}{2},\frac{q}{2}}^{m,n}(\mathbb{R})\right)$$

$$= \pi^{mn} \frac{\Gamma\left(\frac{2}{p} + 1\right)^{mn} \Gamma\left(\frac{2m}{q} + 1\right)^n}{\Gamma\left(\frac{2nm}{q} + 1\right) \Gamma\left(\frac{2m}{p} + 1\right)^n}.$$

Proof. We start by using the substitution $u_{j,k} = r_{j,k} \cos\varphi_{j,k}$ and $v_{j,k} = r_{j,k} \sin\varphi_{j,k}$, then

$$\mathrm{vol}(B_{p,q}^{m,n}(\mathbb{C})) = \int_{B_{p,q}^{m,n}(\mathbb{C})} 1\, d(u, v) = (2\pi)^{mn} \int_{\left\{ r \ge 0 : \sum_{j=1}^n \left(\sum_{k=1}^m r_{j,k}^p \right)^{q/p} \le 1 \right\}} \prod_{j=1}^n \prod_{k=1}^m r_{j,k}\, dr$$

and using $s_{j,k} = r_{j,k}^p$ we obtain with $A_{p,q}^{m,n}(t) = \left\{ s \in \mathbb{R}^{n \times m} : s \geq 0 \text{ and } \sum_{j=1}^{n} \left(\sum_{k=1}^{m} s_{j,k} \right)^{q/p} \leq t \right\}$

$$= \pi^{mn} \left(\frac{2}{p} \right)^{mn} \int_{s \in A_{\frac{p}{2},\frac{q}{2}}^{m,n}(1)} \prod_{j=1}^{n} \prod_{k=1}^{m} s_{j,k}^{\frac{2}{p}-1} ds.$$

Now comparing the last integral to (6) immediately implies

$$\text{vol}(B_{p,q}^{m,n}(\mathbb{C})) = \left(\frac{\pi}{2} \right)^{mn} \text{vol}\left(B_{\frac{p}{2},\frac{q}{2}}^{m,n}(\mathbb{R}) \right).$$

$\square$

3.4 Special cases

In this section we present our main result Theorem 3 in special cases and compare it with known results. First of all we give a table combining the cases $p, q \in \{1, 2, \infty\}$.

$\text{vol}(B_{p,q}^{m,n}(\mathbb{R}))$	$p = 1$	$p = 2$	$p = \infty$
$q = 1$	$\frac{2^{mn}}{(mn)!}$	$\frac{\pi^{\frac{mn}{2}}}{(mn)!} \frac{(m!)^n}{\Gamma(m/2+1)^n}$	$2^{mn} \frac{(m!)^n}{(mn)!}$
$q = 2$	$\frac{2^{mn}}{(m!)^n} \frac{\Gamma(m/2+1)^n}{\Gamma(mn/2+1)}$	$\frac{\pi^{\frac{mn}{2}}}{\Gamma(mn/2+1)}$	$2^{mn} \frac{\Gamma(m/2+1)^n}{\Gamma(mn/2+1)}$
$q = \infty$	$\frac{2^{mn}}{(m!)^n}$	$\frac{\pi^{\frac{mn}{2}}}{\Gamma(m/2+1)^n}$	2^{mn}

$\underline{q = \infty}$: Now, we consider the special case of $q = \infty$, i.e.

$$B_{p,\infty}^{m,n}(\mathbb{R}) = \left\{ x \in \mathbb{R}^{n \times m} : \max_{1 \leq j \leq n} \sum_{k=1}^{m} |x_{j,k}|^p \leq 1 \right\}$$

which can be reformulated by

$$= \left\{ x \in \mathbb{R}^{n \times m} : \sum_{k=1}^{m} |x_{j,k}|^p \leq 1 \text{ for all } 1 \leq j \leq n \right\}.$$

By using Proposition 2 we can easily calculate

$$\text{vol}(B_{p,\infty}^{m,n}(\mathbb{R})) = \int_{x \in B_{p,\infty}^{m,n}(\mathbb{R})} 1 dx = \left(\int_{\{x \in \mathbb{R}^m : \|x\|_p \leq 1\}} 1 dx \right)^n$$

$$= \left(\text{vol}(B_p^m(\mathbb{R})) \right)^n = 2^{mn} \frac{\Gamma(1/p+1)^{mn}}{\Gamma(m/p+1)^n},$$

11

which corresponds to (5) for $q = \infty$.

$\underline{p = \infty}$: Considering the case $p = \infty$ we can write the unit ball

$$B_{\infty,q}^{m,n}(\mathbb{R}) = \left\{ x \in \mathbb{R}^{n \times m} : \sum_{j=1}^{n} \max_{1 \leq k \leq m} |x_{j,k}|^q \leq 1 \right\}$$

and we get

$$\mathrm{vol}(B_{\infty,q}^{m,n}(\mathbb{R})) = 2^{mn} \int\limits_{x \geq 0, x \in B_{\infty,q}^{m,n}(\mathbb{R})} 1 \, dx = 2^{mn} \int\limits_{\substack{\lambda \geq 0 \\ 0 \leq \lambda_1^q + \cdots + \lambda_n^q \leq 1}} \int\limits_{\substack{x \geq 0 \\ \max_{k=1,\ldots,m} x_{j,k} = \lambda_j}} 1 \, dx \, d\lambda$$

$$= 2^{mn} \int\limits_{\substack{\lambda \geq 0 \\ 0 \leq \lambda_1^q + \cdots + \lambda_n^q \leq 1}} \prod_{j=1}^{n} (m \lambda_j^{m-1}) \, d\lambda = 2^{mn} m^n \int\limits_{\substack{\lambda \geq 0 \\ 0 \leq \lambda_1^q + \cdots + \lambda_n^q \leq 1}} \prod_{j=1}^{n} \lambda_j^{m-1} \, d\lambda,$$

where the last but one identity follows from the fact that the surface of the set $\{y \in \mathbb{R}^m : y \geq 0 \text{ and } \max_{k=1,\ldots,m} y_k = \lambda\}$ is equal to $m\lambda^{m-1}$.

We put $\tau_j = \lambda_j^q$ and get

$$\mathrm{vol}(B_{\infty,q}^{m,n}) = 2^{mn} m^n \int\limits_{\substack{\tau \geq 0 \\ 0 \leq \tau_1 + \cdots + \tau_n \leq 1}} \prod_{j=1}^{n} \left(\tau_j^{(m-1)/q} \cdot \frac{1}{q \tau_j^{1-1/q}} \right) d\tau$$

$$= 2^{mn} \frac{m^n}{q^n} \int\limits_{\tau \in A_{\frac{p}{m}, \frac{q}{m}}^{1,n}(1)} \prod_{j=1}^{n} \tau_j^{m/q - 1} \, d\tau.$$

Comparing the last integral to (6) we see the connection to $\mathrm{vol}(B_{q/m}^n(\mathbb{R}))$ and we obtain

$$\mathrm{vol}(B_{\infty,q}^{m,n}) = 2^{n(m-1)} \, \mathrm{vol}(B_{q/m}^n(\mathbb{R})) = 2^{mn} \frac{\Gamma(m/q + 1)^n}{\Gamma(nm/q + 1)},$$

i.e. (5) for $p = \infty$.

$\underline{p = q}$: In this case we can write

$$B_{p,p}^{m,n}(\mathbb{R}) = \left\{ x \in \mathbb{R}^{n \times m} : \sum_{j=1}^{n} \sum_{k=1}^{m} |x_{j,k}|^p \leq 1 \right\}$$

$$= B_p^{mn}(\mathbb{R})$$

and we easily obtain from Proposition 2 the expected result

$$\mathrm{vol}(B_{p,p}^{m,n}(\mathbb{R})) = 2^{mn} \frac{\Gamma(1/p + 1)^{mn}}{\Gamma(mn/p + 1)}.$$

$\underline{m = 1 \text{ or } n = 1}$**:**   In these cases we directly see

$$B_{p,q}^{1,n}(\mathbb{R}) = B_q^n(\mathbb{R}) \qquad \text{as well as} \qquad B_{p,q}^{m,1}(\mathbb{R}) = B_p^m(\mathbb{R})$$

and Proposition 2 then coincides with (5) in Theorem 3.

4   Closing remarks

The elements of the space $\mathbb{R}^{n \times m}$ are usually identified with $n \times m$ matrices. The most important matrix norms (like the nuclear norm or the spectral norm) are, however, not given by simple conditions on the entries of the matrix but rather on its singular values. We refer to [18] for results on volumes of unit balls with respect to these norms.

Theorem 3 gives the volume of the unit ball of $\ell_q^n(\ell_p^m)$. On the other hand, it leaves a number of questions open. Let us briefly discuss the most important ones.

- On many occurrences in mathematics it is more convenient to work with Lorentz sequence spaces instead of Lebesgue sequence spaces. If $0 < p \leq \infty$, the weak-ℓ_p^m (quasi-)norm is defined as

$$\|x\|_{p,\infty} = \max_{j=1,\dots,m} j^{1/p} x_j^*,$$

  where $(x_j^*)_{j=1}^m$ is the non-increasing rearrangement of $x \in \mathbb{R}^m$. To our best knowledge, no closed formula for the volume of the unit ball $\{x \in \mathbb{R}^m : \|x\|_{p,\infty} \leq 1\}$ is available in the literature. The same holds true for the general Lorentz sequence spaces $\ell_{p,q}^m$ with $0 < q < \infty$ if $q \neq p$.

- Comparing (5) with (4), we observe that

$$\mathrm{vol}(B_{p,q}^{m,n}(\mathbb{R})) = \mathrm{vol}(B_q^{n \cdot m}(\mathbb{R})) \cdot \frac{[\mathrm{vol}(B_p^m(\mathbb{R}))]^n}{[\mathrm{vol}(B_q^m(\mathbb{R}))]^n}. \tag{19}$$

  This formula appeared implicitly already in (14). We leave it open if this formula has some geometric or combinatoric interpretation, which would allow for an non-analytical proof of Theorem 3.

References

[1] F. R. Bach, *Consistency of the group lasso and multiple kernel learning*, J. Mach. Learn. Res. **9** (2008), 1179–1225.

[2] F. Barthe, M. Csörnyei, A. Naor, *A note on simultaneous polar and Cartesian decomposition*, in: Geometric Aspects of Functional Analysis, Lecture Notes in Mathematics, Springer, Berlin, 2003.

[3] B. Carl, *Entropy numbers, s-numbers, and eigenvalue problems*, J. Funct. Anal. 41 (1981), no. 3, 290–306.

[4] B. Carl, H. Triebel, *Inequalities between eigenvalues, entropy numbers, and related quantities of compact operators in Banach spaces*. Math. Ann. **251** (1980), no. 2, 129–133.

[5] E. J. Candés, J. Romberg, T. Tao, *Stable signal recovery from incomplete and inaccurate measurements*, Comm. Pure Appl. Math. **59** (2006), 1207–1223.

[6] E. J. Candés, T. Tao, *Near-optimal signal recovery from random projections: universal encoding strategies?*, IEEE Trans. Inform. Theory **52** (2006), 5406–5425.

[7] I. A. Davidovich, *Widths of function classes and sets of finite volume in spaces with mixed norms*, Dissertation, Moscow, 1995.

[8] P.J. Davis, *6. Gamma function and related functions*, in M. Abramowitz, I.A. Stegun, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, New York: Dover Publications, 1972.

[9] J.P.G.L. Dirichlet, *Sur une nouvelle méthode pour la détermination des intégrales multiples*, Journal de Mathématiques Pures et Appliquées **4** (1839), 164–168.

[10] D. L. Donoho, *Compressed sensing*, IEEE Trans. Inform. Theory **52** (2006), 1289–1306.

[11] D. E. Edmunds, H. Triebel, *Function spaces, entropy numbers, differential operators*, Cambridge Tracts in Math. **120**, 1996.

[12] D. E. Edmunds, Yu. Netrusov, *Schütts theorem for vector-valued sequence spaces*, J. Appr. Theory **178** (2014), 13–21.

[13] J. Edwards, *A Treatise on the Integral Calculus*, Vol. II, Chelsea Publishing Company, New York, 1922.

[14] L. Markhasin, *Discrepancy and integration in function spaces with dominating mixed smoothness*, Friedrich-Schiller-Universität Jena, Dissertation, 2012.

[15] S. Mayer, S. Dirksen, T. Ullrich, *Gelfand and entropy numbers of $\ell_p(\ell_q)$ balls*, in preparation.

[16] L. Meier, S. van de Geer, P. Bühlmann, *The group Lasso for logistic regression*, J. R. Stat. Soc. B **70** (2008), no. 1, 53–71.

[17] G. Pisier, *The volume of convex bodies and Banach space geometry*, Cambridge Tracts in Math. **94**, 1989.

[18] J. Saint-Raymond, *Le volume des idéaux d'opérateurs classiques*, Studia Math. **80** (1984), no. 1, 63–75.

[19] G. Schechtman and J. Zinn, *On the volume of the intersection of two L_p^n balls*, Proc. AMS **110** (1990), (1), 217–224.

[20] R. Tibshirani, *Regression shrinkage and selection via the Lasso*, J. R. Stat. Soc. B **58** (1996), 267–288.

[21] A.A. Vasil'eva, *Kolmogorov and linear widths of the weighted Besov classes with singularity at the origin*, J. Appr. Theory **167** (2013), 1–41.

[22] J. Vybíral, *Function spaces with dominating mixed smoothness*, Diss. Math. **436** (2006), 1–73.

[23] J. Vybíral, *Widths of embeddings in function spaces*, J. Compl. **24** (2008), 545–570.

[24] M. Yuan, Y. Lin, *Model selection and estimation in regression with grouped variables*, J. R. Stat. Soc. B **68** (2006), (1), 49–67.