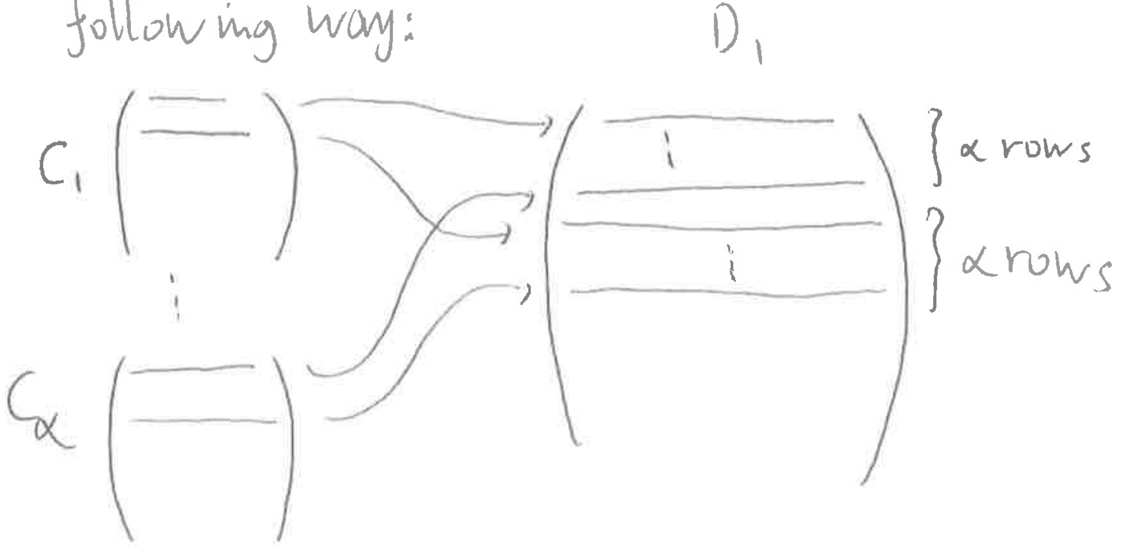
(1.4.8) Construction of Order α digital nets

Let $C_1, \ldots, C_{\alpha s} \in \mathbb{Z}_b^{m \times m}$ be the generating matrices of a digital $(t, m, \alpha s)$-net.

We define the matrices $D_1, \ldots, D_s \in \mathbb{Z}_b^{\alpha m \times m}$ in the following way:



and so on. The matrix D_j is obtained by interlacing the matrices $C_{\alpha j}, C_{\alpha j+1}, \ldots, C_{\alpha(j+1)-1}$.

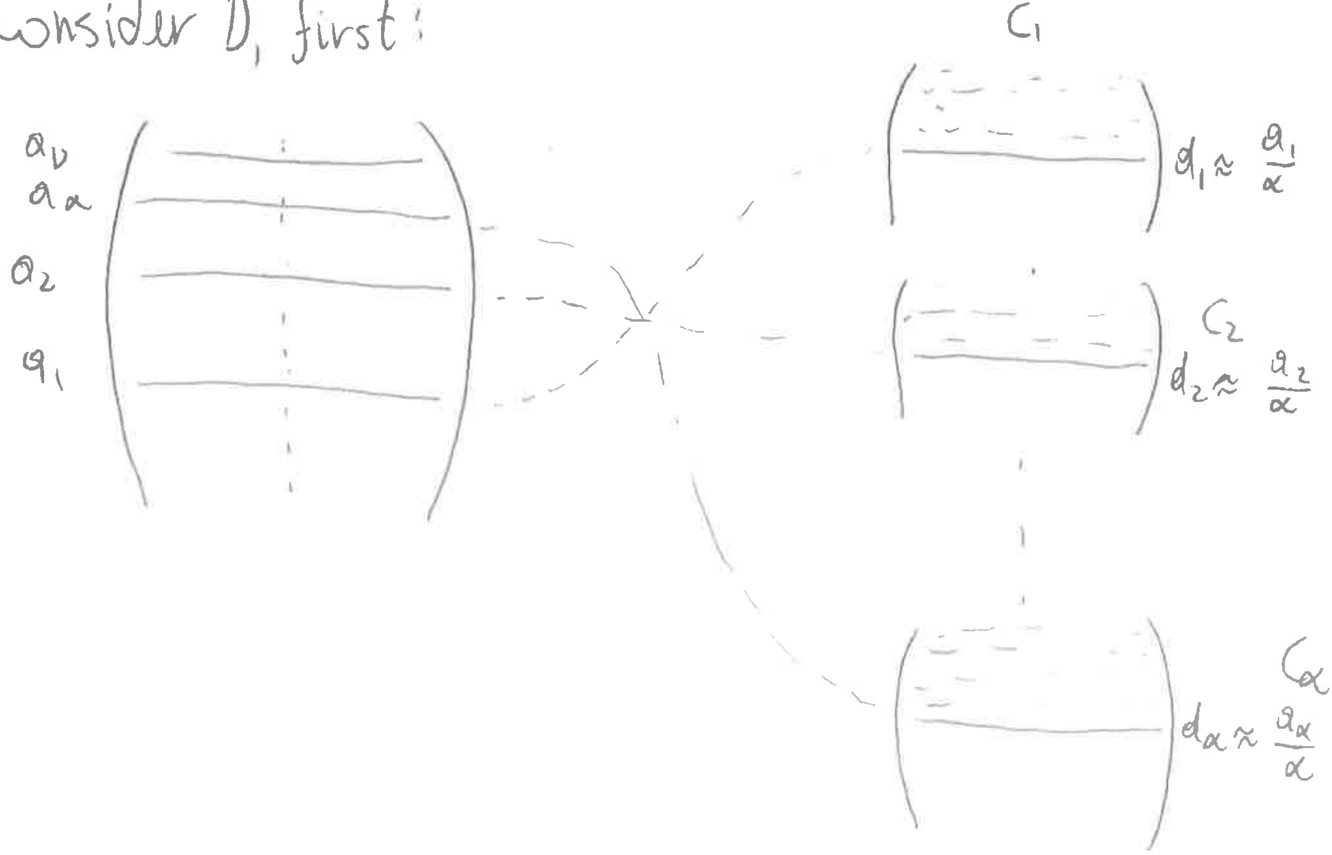
The number of matrices we interlace (in this case α) to get one new matrix is called the interlacing factor.

Let $C_1, \dots, C_{s\alpha}$ be the generating matrices of a digital $(t, m, \alpha s)$-net and let $D_1, \dots, D_s \in \mathbb{Z}_\ell^{\alpha m \times m}$ be the matrices obtained by interlacing $C_1, \dots, C_{s\alpha}$.

We check the linear independence condition for order α digital (t_α, m, s)-nets. Let $a_{j1} > a_{j2} > \dots > a_{j\nu_j} > 0$ such that

$$\sum_{j=1}^{s} \left(a_{j1} + \dots + a_{j\min(\alpha,\nu_j)} \right) \leq \alpha m - t_\alpha .$$

Consider D_1 first:



$$\sum_{j=1}^{s} \left(a_{j,1} + \dots + a_{j,\min(\alpha,\nu_j)} \right) \approx \alpha \sum_{j=1}^{s\alpha} d_j \le \alpha(m-t)$$

$\Rightarrow$ Choose $t_\alpha \approx \alpha t$.

For more information see

J. Dick and F. Pillichshammer, Digital Nets and Sequences. Discrepancy Theory and Quasi-Monte Carlo Integration. Cambridge University Press, Cambridge, 2010.

(2) Point distributions on the sphere

(2.1) Stolarsky's invariance principle

We consider the sphere

$$\mathbb{S}^d = \left\{ \underline{z} = (z_1, z_2, \dots, z_{d+1}) \in \mathbb{R}^{d+1} : z_1^2 + z_2^2 + \dots + z_{d+1}^2 = 1 \right\}.$$

We consider numerical integration on $\mathbb{S}^d$

$$\int_{\mathbb{S}^d} f(\underline{x}) \, d\sigma_d(\underline{x}) \simeq \frac{1}{N} \sum_{n=0}^{N-1} f(\underline{x}_n)$$

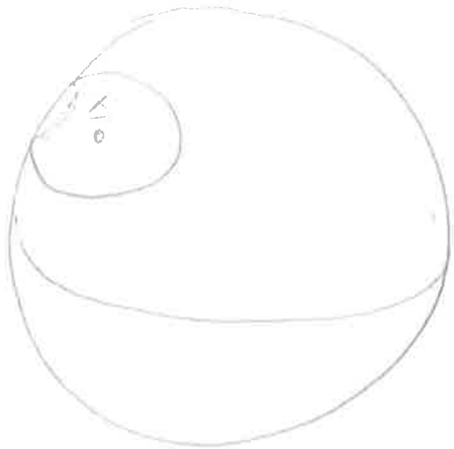
where σ_d is the normalized Lebesgue surface area measure (i.e. $\int_{\mathbb{S}^d} 1 \, d\sigma_d(\underline{x}) = 1$).

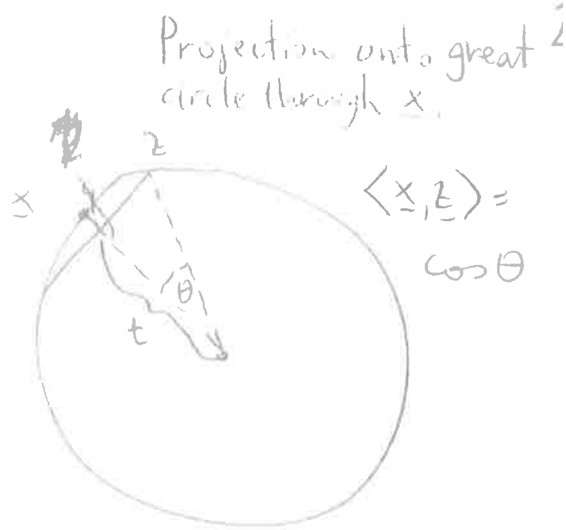
In the following we relate the integration error in a certain reproducing kernel Hilbert space to the spherical cap discrepancy.

A spherical cap centered at $\underline{x} \in \mathbb{S}^d$ with height $t \in [-1, 1]$ is the set

$$C(\underline{x}, t) = \left\{ \underline{z} \in \mathbb{S}^d : \underbrace{\langle \underline{x}, \underline{z} \rangle}_{= x_1 z_1 + \dots + x_{d+1} z_{d+1}} \geq t \right\}$$

$\mathbb{S}^2$





In particular $C(\underline{x}, 1) = \{\underline{x}\}$ and $C(\underline{x}, -1) = \mathbb{S}^d$.

For a measurable set $J \subseteq \mathbb{S}^d$ let $\sigma_d(J) = \int_{\mathbb{S}^d} 1_J(\underline{x})\, d\sigma_d(\underline{x})$ be the measure of J.

Let $P = \{\underline{x}_0, \underline{x}_1, \ldots, \underline{x}_{N-1}\} \subseteq \mathbb{S}^d$ be a point set on the sphere.

We define the spherical cap L_2-discrepancy by

$$L_2(P) = \left(\int_{-1}^{1} \int_{\mathbb{S}^d} \left| \sigma_d(C(\underline{z}, t)) - \frac{1}{N} \sum_{n=0}^{N-1} 1_{C(\underline{z}, t)}(\underline{x}_n) \right|^2 d\sigma_d(\underline{z})\, dt \right)^{1/2}$$

Another criterion which can be applied to points on the sphere is the sum of distances

$$\frac{1}{N^2} \sum_{n,m=0}^{N-1} \| \underline{x}_n - \underline{x}_m \|,$$

where $\|\cdot\|$ is just the Euclidean distance. The aim is to construct point sets $\{\underline{x}_0, \underline{x}_1, \ldots, \underline{x}_{N-1}\} \subseteq \mathbb{S}^d$ such that the sum of distances is as large as possible. Think of N points repelling each other to make the sum as large as possible.

A reproducing kernel Hilbert space on the sphere.

For $\underline{x}, \underline{y} \in \mathbb{S}^d$ let

$$K : \mathbb{S}^d \times \mathbb{S}^d \to \mathbb{R}$$

$$K(\underline{x}, \underline{y}) = \int_{-1}^{1} \int_{\mathbb{S}^d} 1_{C(\underline{z};t)}(\underline{x})\, 1_{C(\underline{z};t)}(\underline{y})\, d\sigma_d(\underline{z})\, dt .$$

$$= \int_{-1}^{1} \int_{\mathbb{S}^d} 1_{C(\underline{x};t)}(\underline{z})\, 1_{C(\underline{y};t)}(\underline{z})\, d\sigma_d(\underline{z})\, dt .$$

The corresponding r.k.H.s consists of all functions of the form

$$f_i(\underline{x}) = \int_{-1}^{1} \int_{\mathbb{S}^d} g_i(\underline{z};t)\, 1_{C(\underline{z};t)}(\underline{x})\, d\sigma_d(\underline{z})\, dt ,$$

$$g(\underline{z};t) \in L_2(\mathbb{S}^d \times [-1,1]).$$

The inner product is given by

$$\langle f_1, f_2 \rangle = \int_{-1}^{1} \int_{\mathbb{S}^d} g_1(\underline{z};t)\, g_2(\underline{z};t)\, d\sigma_d(\underline{z})\, dt .$$

and the corresponding norm by

$$\|f\| = \left(\int_{-1}^{1} \int_{\mathbb{S}^d} |g(\underline{z};t)|^2\, d\sigma_d(\underline{z})\, dt \right)^{1/2}$$

We have

$$\|K(\cdot,y)\|^2 = \int_{-1}^{1}\int_{\mathbb{S}^d} 1^2_{C(\underline{z};t)}(y)\, d\sigma_d(z)\, dt < \infty$$

Hence $K \in \mathcal{H}$ and

$$\langle f, K(\cdot,y)\rangle = \int_{-1}^{1}\int_{\mathbb{S}^d} g(\underline{z};t)\, 1_{C(\underline{z};t)}(y)\, d\sigma_d(\underline{z})\, dt = f(y),$$

therefore K is the reproducing kernel of $\mathcal{H}$.

There is an explicit formula for the reproducing kernel:

$$K(\underline{x},\underline{y}) = \int_{-1}^{1}\int_{\mathbb{S}^d} 1_{C(\underline{z},t)}(\underline{x})\, 1_{C(\underline{z};t)}(\underline{y})\, d\sigma_d(\underline{z})\, dt$$

$$= \int_{\mathbb{S}^d}\int_{-1}^{1} 1_{C(\underline{z};t)}(\underline{x})\, 1_{C(\underline{z};t)}(y)\, dt\, d\sigma_d(z)$$

$$= \int_{\mathbb{S}^d}\int_{-1}^{1} 1_{C(\underline{x};t)}(z)\, 1_{C(\underline{y},t)}(z)\, dt\, d\sigma_d(z)$$

$$C(\underline{x};t) = \{\underline{z} \in \mathbb{S}^d : \langle \underline{z}, \underline{x}\rangle \geq t\}$$
$$C(\underline{y};t) = \{\underline{z} \in \mathbb{S}^d : \langle \underline{z}, \underline{y}\rangle \geq t\}$$

$$= \int_{\mathbb{S}^d}\int_{-1}^{1} 1_{\langle \underline{z},\underline{x}\rangle \geq t}\, 1_{\langle \underline{z},\underline{y}\rangle \geq t}\, dt\, d\sigma_d(z)$$

$$K(\underline{x},\underline{y}) = \int_{\mathbb{S}^d} \int_{-1}^{\min(\langle \underline{z},\underline{x}\rangle,\langle \underline{z},\underline{y}\rangle)} dt\, d\sigma_d(\underline{z})$$

$$= \int_{\mathbb{S}^d} \left[1+\min(\langle \underline{z},\underline{x}\rangle,\langle \underline{z},\underline{y}\rangle)\right] d\sigma_d(\underline{z})$$

$$= 1 + \int_{\mathbb{S}^d} \min(\langle \underline{z},\underline{x}\rangle,\langle \underline{z},\underline{y}\rangle)\, d\sigma_d(\underline{z})$$

We have

$$\min(\langle \underline{z},\underline{x}\rangle,\langle \underline{z},\underline{y}\rangle) = \frac{1}{2}\left[\langle \underline{z},\underline{x}\rangle + \langle \underline{z},\underline{y}\rangle - |\langle \underline{z},\underline{x}-\underline{y}\rangle|\right]$$

and

$$\int_{\mathbb{S}^d} \langle \underline{z},\underline{x}\rangle\, d\sigma_d(\underline{z}) = \int_{\mathbb{S}^d} \langle \underline{z},\underline{y}\rangle\, d\sigma_d(\underline{z}) = 0$$

by symmetry

Thus

$$K(\underline{x},\underline{y}) = 1 - \frac{1}{2}\int_{\mathbb{S}^d} |\langle \underline{z},\underline{x}-\underline{y}\rangle|\, d\sigma_d(\underline{z})$$

$$= 1 - \frac{1}{2}\|\underline{x}-\underline{y}\| \int_{\mathbb{S}^d} \left|\left\langle \underline{z}, \frac{\underline{x}-\underline{y}}{\|\underline{x}-\underline{y}\|}\right\rangle\right| d\sigma_d(\underline{z})$$

$$= 1 - \|\underline{x}-\underline{y}\| \underbrace{\frac{1}{2}\int_{\mathbb{S}^d} |\langle \underline{z},\underline{p}\rangle|\, d\sigma_d(\underline{z})}_{=: C_d}, \quad \underline{p}\in\mathbb{S}^d$$

We have

$$C_d = \frac{1}{2}\int_{\mathbb{S}^d} |\langle \underline{z}, \underline{p}\rangle| \, d\sigma_d(\underline{z}) = \frac{1}{d}\frac{\Gamma\left(\frac{d+1}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{d}{2}\right)}$$

Therefore

$$K(\underline{x},\underline{y}) = 1 - C_d \|\underline{x}-\underline{y}\|.$$

The worst-case integration error is given by

$$e^2(P,\mathcal{H}) = \int_{\mathbb{S}^d}\int_{\mathbb{S}^d} K(\underline{x},\underline{y})\, d\sigma_d(x)\, d\sigma_d(y) - \frac{2}{N}\sum_{n=0}^{N-1}\int_{\mathbb{S}^d} K(\underline{x}_n, x)\, d\sigma_d(x)$$

$$+ \frac{1}{N^2}\sum_{n,m=0}^{N-1} K(\underline{x}_n, x_m).$$

We have

$$\int_{\mathbb{S}^d} K(\underline{y},\underline{x})\, d\sigma_d(\underline{x}) = \int_{-1}^{1}\int_{\mathbb{S}^d}\int_{\mathbb{S}^d} 1_{C(\underline{x};t)}(\underline{z})\, 1_{C(\underline{y},t)}(\underline{z})\, d\sigma_d(x)\, d\sigma_d(\underline{z})\, dt$$

$$= \int_{-1}^{1}\int_{\mathbb{S}^d}\underbrace{\int_{\mathbb{S}^d} 1_{C(\underline{z};t)}(\underline{x})\, d\sigma_d(\underline{x})}_{\sigma_d(C(\underline{z};t))}\, 1_{C(\underline{y};t)}(\underline{z})\, d\sigma_d(\underline{z})\, dt$$

*

$$= \int_{-1}^{1} \sigma_d(C(\underline{p};t)) \int_{\mathbb{S}^d} 1_{C(\underline{y};t)}(\underline{z})\, d\sigma_d(\underline{z})\, dt$$

$$= \int_{-1}^{1} \left(\sigma_d \left(C(p, t) \right) \right)^2 dt .$$

In another way:

$$\int_{\mathbb{S}^d} K(x, y)\, d\sigma_d(y) = 1 - C_d \int_{\mathbb{S}^d} \|x - y\|\, d\sigma_d(y)$$

$$= 1 - C_d \int_{\mathbb{S}^d} \|p - y\|\, d\sigma_d(y) ,$$

since the Lebesgue surface measure is rotationally invariant.

Thus

$$e^2(P, \mathcal{H}) = 1 - C_d \int_{\mathbb{S}^d}\int_{\mathbb{S}^d} \|x - y\|\, d\sigma_d(x)\, d\sigma_d(y)$$

$$-2 + C_d\, 2 \int_{\mathbb{S}^d} \|p - y\|\, d\sigma_d(y)$$

$$+1 - C_d \frac{1}{N^2} \sum_{n,m=0}^{N-1} \|x_n - x_m\|$$

$$= C_d \left[\int_{\mathbb{S}^d}\int_{\mathbb{S}^d} \|x - y\|\, d\sigma_d(x)\, d\sigma_d(y) - \frac{1}{N^2} \sum_{n,m=0}^{N-1} \|x_n - x_m\| \right] \tag{2.19}$$

Maximizing $\frac{1}{N^2} \sum_{n,m=0}^{N-1} \|x_n - x_m\|$ minimizes the worst-case error.

On the other hand we can relate the worst-case error to the spherical cap L_2-discrepancy. We have

$$f(\underline{x}) = \int_{-1}^{1} \int_{\mathbb{S}^d} g(\underline{z};t)\, 1_{C(\underline{z};t)}(\underline{x})\, d\sigma_d(\underline{z})\, dt.$$

Thus

$$\frac{1}{N}\sum_{n=0}^{N-1} f(\underline{x}_n) = \int_{-1}^{1}\int_{\mathbb{S}^d} g(\underline{z};t)\, \frac{1}{N}\sum_{n=0}^{N-1} 1_{C(\underline{z};t)}(\underline{x}_n)\, d\sigma_d(\underline{z})\, dt$$

and

$$\int_{\mathbb{S}^d} f(\underline{x})\, d\sigma_d(\underline{x}) = \int_{-1}^{1}\int_{\mathbb{S}^d} g(\underline{z};t) \int_{\mathbb{S}^d} 1_{C(\underline{z};t)}(\underline{x})\, d\sigma_d(\underline{x})\, d\sigma_d(\underline{z})\, dt$$

$$= \int_{-1}^{1}\int_{\mathbb{S}^d} g(\underline{z};t)\, \sigma_d(C(\underline{z};t))\, d\sigma_d(\underline{z})\, dt.$$

The integration error can therefore be bounded in the following way:

$$e(f) = \int_{\mathbb{S}^d} f(\underline{x})\, d\sigma_d(\underline{x}) - \frac{1}{N}\sum_{n=0}^{N-1} f(\underline{x}_n) = \int_{-1}^{1}\int_{\mathbb{S}^d} g(\underline{z};t) \left[\sigma_d(C(\underline{z};t)) - \frac{1}{N}\sum_{n=0}^{N-1} 1_{C(\underline{z};t)}(\underline{x}_n)\right] d\sigma_d(\underline{z})\, dt$$

and by using Hölder's inequality we obtain

$$|e(g)| \leq \left(\int_{-1}^{1} \int_{\mathbb{S}^d} |g(\underline{z};t)|^2 \, d\sigma_d(z)\, dt \right)^{1/2} \left(\int_{-1}^{1} \int_{\mathbb{S}^d} \left| \sigma_d(C(\underline{z};t)) - \frac{1}{N} \sum_{n=0}^{N-1} 1_{C(\underline{z};t)}(x_n) \right|^2 d\sigma_d(z)\, dt \right)^{1/2}$$

$$= \|g\| \; L_2(P).$$

By setting

$$g(\underline{z};t) = \frac{\Delta_P(\underline{z};t)}{\|\Delta_P(\underline{z};t)}$$

where

$$\Delta_P(\underline{z};t) = \sigma_d(C(\underline{z};t)) - \frac{1}{N} \sum_{n=0}^{N-1} 1_{C(\underline{z};t)}(\underline{x}_n)$$

is the local discrepancy function, we obtain

$$e(P, \mathcal{H}) = \sup_{\|g\| \leq 1} |e(g)| = L_2(P).$$

Using (2.1a) we obtain

$$\int_{-1}^{1} \int_{\mathbb{S}^d} \left| \sigma_d(C(\underline{z};t)) - \frac{1}{N} \sum_{n=0}^{N-1} 1_{C(\underline{z};t)}(\underline{x}_n) \right|^2 d\sigma_d(\underline{z})\, dt$$

$$= C_d \left[\int_{\mathbb{S}^d} \int_{\mathbb{S}^d} \|\underline{x} - \underline{y}\| \, d\sigma_d(\underline{x})\, d\sigma_d(y) - \frac{1}{N^2} \sum_{n,m=0}^{N-1} \|\underline{x}_n - \underline{x}_m\| \right].$$

Stolarsky's invariance principle is now

$$\frac{1}{N^2}\sum_{n,m=0}^{N-1} \|\underline{x}_n - \underline{x}_m\| + \frac{1}{C_d}\int_{-1}^{1}\int_{\mathbb{S}^d} \left|\sigma_d(C(\underline{z},t)) - \frac{1}{N}\sum_{n=0}^{N-1} 1_{C(\underline{z},t)}(\underline{x}_n)\right|^2 d\sigma_d(\underline{z})\,dt$$

$$= \int_{\mathbb{S}^d}\int_{\mathbb{S}^d} \|\underline{x}-\underline{y}\|\, d\sigma_d(\underline{x})\, d\sigma_d(\underline{y}) .$$

Thus the sum of distances plus $\frac{1}{C_d}$ times the ~~L2~~ spherical cap L_2 discrepancy are invariant of the choice of point set P.

(2.2) Spherical harmonics

We consider only the sphere $\mathbb{S}^2$. Just like Fourier series on the interval $[0,2\pi)$, we have series expansions on the sphere $\mathbb{S}^2$ based on spherical harmonics.

On $[0,2\pi)$ we write

$$f(x) = \sum_{k\in\mathbb{Z}} \hat{f}(k)\, e^{ikx}.$$

If we use the circle $\mathbb{S}^1 = \{\underline{x} \in \mathbb{R} : x_1^2 + x_2^2 = 1\}$, then using the point $(\cos\theta, \sin\theta)$ with $0 \leq \theta < 2\pi$, we write

$$f(\cos\theta, \sin\theta) = \sum_{k\in\mathbb{Z}} \hat{f}(k)\, e^{i\theta k}.$$

The functions e^{ikx} are L_2 orthogonal.

On $\mathbb{S}^2$ we also have a set of orthogonal functions $Y_{\ell,k}$, where $\ell \in \mathbb{N}_0$ and $k \in \mathbb{Z}$ with $|k| \leq \ell$. We have

$$Y_{0,0} = 1$$

$$Y_{1,-1} = \sqrt{\tfrac{3}{2}}\,(x - iy)$$

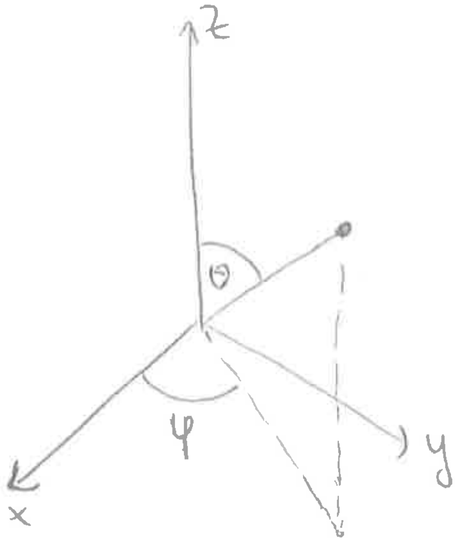
$$Y_{1,0} = \sqrt{\tfrac{3}{2}}\,z$$

$$Y_{1,1} = \sqrt{\tfrac{3}{2}}\,(x + iy)$$

$$\vdots$$

These are homogeneous polynomials (each monomial has the same total degree).

It is convenient to use spherical coordinates $(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$ where $0 \le \varphi < 2\pi$ and $0 \le \theta \le \pi$.



Then we can write the spherical harmonics as

$$Y_{\ell,k}(\varphi,\theta) = e^{ik\varphi} P_\ell^k(\cos\theta)\sqrt{(2\ell+1)\frac{(\ell-k)!}{(\ell+k)!}}$$

where P_ℓ^k is the associated Legendre polynomial

$$P_\ell^k(z) = \frac{(-1)^k}{2^\ell \ell!}(1-x^2)^{k/2}\frac{d^{\ell+k}}{dz^{\ell+k}}\left\{(z^2-1)^\ell\right\}$$

For fixed k they satisfy the orthogonality condition

$$\int_{-1}^{1} P_\ell^k(z)\,P_m^k(z)\,dz = \delta_{\ell m}\frac{2(\ell+k)!}{(2\ell+1)(\ell-k)!}$$

and for fixed ℓ they satisfy

$$\int_{-1}^{1}\frac{P_\ell^k(z)\,P_\ell^m(z)}{1-z^2}\,dz = \begin{cases} 0 & \text{if } k \neq m, \\ \dfrac{(\ell+m)!}{m(\ell-m)!} & \text{if } k = m \neq 0, \\ \infty & \text{if } k = m = 0. \end{cases}$$

We have

$$\int_{S^2} Y_{\ell,k} \overline{Y_{\ell',k'}}\, d\sigma_2 = \int_0^{2\pi}\int_0^{\pi} e^{ik\varphi} P_\ell^k(\cos\theta) e^{-ik'\varphi} P_{\ell'}^{k'}(\cos\theta) \sin\theta\, d\theta\, d\varphi$$

$$= \underbrace{\int_0^{2\pi} e^{i(k-k')\varphi}\, d\varphi}_{= \begin{cases} 2\pi & \text{if } k=k' \\ 0 & \text{if } k\neq k' \end{cases}} \underbrace{\int_0^{\pi} P_\ell^k(\cos\theta) P_{\ell'}^{k'}(\cos\theta) \sin\theta\, d\theta}$$

$$z = \cos\theta$$

$$\frac{dz}{d\theta} = -\sin\theta$$

$$= \int_{-1}^{1} P_\ell^k(z) P_{\ell'}^{k'}(z)\, dz$$

Assume $k=k'$:

$$\int_{-1}^{1} P_\ell^k(z) P_{\ell'}^k(z)\, dz = \begin{cases} 0 & \text{if } \ell\neq\ell' \\ \dfrac{2(\ell+k)!}{(2\ell+1)(\ell-k)!} & \text{if } \ell=\ell' \end{cases}$$

Thus: Note $\int_0^{2\pi}\int_0^{\pi} 1 \sin\theta \, d\theta \, d\varphi = 2\pi \left(-\cos\theta\right)\Big|_0^{\pi} = 4\pi$.

Hence

$$\int_{\mathbb{S}^2} Y_{l,k} \, Y_{l',k'} \, d\sigma_2 = 0 \quad \text{if } (l,k) \neq (l',k')$$

and

$$\int_{\mathbb{S}^2} Y_{l,k} Y_{l,k} \, d\sigma_2 = \frac{1}{4\pi}(2l+1)\frac{(l-k)!}{(l+k)!}\int_0^{2\pi}\int_0^{\pi} e^{ik\varphi}\left(P_l^k(\cos\theta)\right)^2 \sin\theta \, d\theta \, d\varphi$$

$$= \frac{1}{2}(2l+1)\frac{(l-k)!}{(l+k)!}\int_{-1}^{1}\left(P_l^k(z)\right)^2 dz = \frac{2l+1}{2}\,\frac{(l-k)!}{(l+k)!}\,\frac{2(l+k)!}{(2l+1)(l-k)!} = 1.$$

The functions $\{Y_{l,k} : l \in \mathbb{N}_0, |k| \leq l\}$ form a complete orthonormal system on $L_2(\mathbb{S}^2)$.

We can define the Fourier series for $f \in L_2(\mathbb{S}^2)$ by

$$f(\underline{x}) = \sum_{l=0}^{\infty} \sum_{k=-l}^{l} \hat{f}(l,k) \, Y_{l,k}(\underline{x}).$$

As for Fourier series, the decay rate of the coefficients $\hat{f}(l,k)$ is associated with the smoothness of the function f.

We can use the decay rate to define a reproducing kernel Hilbert space:

$$K_\alpha(\underline{x},\underline{y}) = \sum_{l=0}^{\infty} \lambda_\alpha(l) \underbrace{\sum_{k=-l}^{l} Y_{l,k}(\underline{x}) \overline{Y_{l,k}(\underline{y})}}_{=\left(\frac{2l+1}{4\pi}\right) P_l(\underline{x}\cdot\underline{y})}$$

where P_l is the Legendre polynomial.

$$\int_{-1}^{1} P_m(z) P_n(z)\,dz = \begin{cases} 0 & \text{if } m \neq n, \\ \frac{2}{2n+1} & \text{if } m = n. \end{cases}$$

Let $f(\underline{x}) = \sum_{l=0}^{\infty} \sum_{k=-l}^{l} \hat{f}(k,l)\, Y_{l,k}(\underline{x})$.

The inner product is given by

$$\langle f, g \rangle = \sum_{l=0}^{\infty} \sum_{k=-l}^{l} \hat{f}(k,l)\, \overline{\hat{g}(k,l)}\, \lambda_\alpha^{-1}(l).$$

Then

$$\langle f, K_\alpha(\cdot,\underline{y}) \rangle = \sum_{l=0}^{\infty} \sum_{k=-l}^{l} \hat{f}(k,l)\, \lambda_\alpha^{-1}(l)\, Y_{l,k}(\underline{y})\, \lambda_\alpha(l)$$

$$= \sum_{l=0}^{\infty} \sum_{k=-l}^{l} \hat{f}(k,l)\, \frac{\lambda_\alpha(l)}{\lambda_\alpha(l)}\, Y_{l,k}(\underline{y}) = f(\underline{y}).$$

For the reproducing kernel

$$K(\underline{x},\underline{y}) = \int_{-1}^{1}\int_{\mathbb{S}^2} 1_{C(\underline{z};t)}(\underline{x})\, 1_{C(\underline{z};t)}(\underline{y})\, d\sigma_2(\underline{z})\, dt$$

we have

$$K(\underline{x},\underline{y}) = \sum_{\ell=0}^{\infty} b_\ell \sum_{k=-\ell}^{\ell} Y_{\ell,k}(\underline{x})\, \overline{Y_{\ell,k}(\underline{y})}$$

with $b_\ell \asymp (1+\ell)^{-3}$. This space is associated with the Sobolev space $\mathcal{H}_s$ with $s = 3/2$.

In general, if we have a reproducing kernel

$$K_\alpha(\underline{x},\underline{y}) = \sum_{\ell=0}^{\infty} \lambda_\alpha(\ell) \sum_{k=-\ell}^{\ell} Y_{\ell k}(\underline{x})\, \overline{Y_{\ell k}(\underline{y})}$$

with $\lambda_\alpha(\ell) \asymp (1+\ell)^{-2\alpha}$,

then the r. k. H. s $\mathcal{H}_\alpha$ is the Sobolev space of smoothness α

(2.3) Spherical designs

A set $\underline{x}_0, \underline{x}_1, \dots, \underline{x}_{N-1}$ is called a spherical t-design if

$$\int_{\mathbb{S}^d} P(\underline{x})\, d\sigma_d(\underline{x}) = \frac{1}{N}\sum_{n=0}^{N-1} P(\underline{x}_n)$$

for all polynomials of total degree at most t.
(Note that for $Y_{\ell,k}$ from the previous section, the $Y_{\ell k}$ are polynomials of total degree ℓ.)

For $d, t \in \mathbb{N}$ let $N(d,t)$ be the minimal number of points in a spherical t-design in $\mathbb{S}^d$.

Bondarenko, Radchenko, Viazovska (2013) showed that there exist constants $C_d > 0$ independent of t, such that

$$N(d,t) \leq C_d\, t^d \text{ for all } t \in \mathbb{N}.$$

Spherical t-designs achieve a convergence rate of the worst-case error of order

$$e(P_t, \mathcal{H}_\alpha) \lesssim t^{-s}.$$

For $N \asymp t^d$ we obtain the optimal rate of convergence

$$e(P_t, \mathcal{H}_\alpha) \lesssim N^{-\frac{s}{d}} \quad \forall s > \tfrac{1}{2}.$$

Obtaining explicit constructions of point sets is a very difficult open problem.

(2.4) Projecting points from the square to $\mathbb{S}^2$

Spherical coordinates are given by

$$U(\theta,\varphi) = (\sin \pi\theta \cos 2\pi\varphi, \sin\pi\theta \sin 2\pi\varphi, \cos\pi\theta)$$

for $0 \le \theta, \varphi \le 1$.

However, this transformation is not preserving area, i.e. for a measurable subset $A \subseteq [0,1]^2$ we do not have

$$\underbrace{\lambda(A)}_{\text{Lebesgue measure}} = \sigma_2(\{U(\theta,\varphi) : (\theta,\varphi) \in A\}).$$

Thus equid distribution in $[0,1]^2$ lifted to the sphere using U does not yield equi-distribution on $\mathbb{S}^2$.

However, the Lambert transform does preserve area

$$T(\theta,\varphi) = \left(2\sqrt{\theta-\theta^2}\cos 2\pi\varphi, 2\sqrt{\theta-\theta^2}\sin 2\pi\varphi, 2\theta - 1\right)$$

$$T : [0,1)^2 \to \mathbb{S}^2.$$

I.e. we have

$$\lambda(A) = \sigma_2(T(A)) \quad \forall \text{measurable } A \subseteq [0,1]^2.$$

Let $P \subseteq [0,1]^2$ be an N-element point set. We consider the spherical cap L_∞ discrepancy of $T(P)$

$$L_\infty(T(P)) = \sup_{\substack{\underline{z} \in \mathbb{S}^d \\ -1 \le t \le 1}} \left| \sigma_2(C(\underline{z};t)) - \frac{1}{N}\sum_{n=0}^{N-1} 1_{C(\underline{z};t)}(T(\underline{x}_n)) \right|,$$

where $P = \{\underline{x}_0, \underline{x}_1, \ldots, \underline{x}_{N-1}\}$.

Since T is area preserving we have

$$\sigma_2(C(\underline{z};t)) - \frac{1}{N}\sum_{n=0}^{N-1} 1_{C(\underline{z};t)}(T(\underline{x}_n))$$

$$= \lambda\left(T^{-1}(C(\underline{z};t))\right) - \frac{1}{N}\sum_{n=0}^{N-1} 1_{T^{-1}(C(\underline{z};t))}(\underline{x}_n).$$

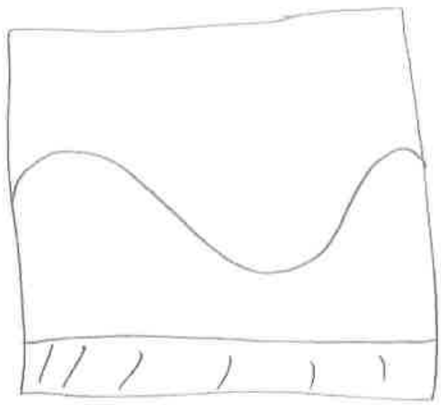
Thus

$$L_\infty(T(P)) = \sup_{\substack{\underline{z} \in \mathbb{S}^2 \\ -1 \le t \le 1}} \left| \lambda\left(T^{-1}(C(\underline{z};t))\right) - \frac{1}{N}\sum_{n=0}^{N-1} 1_{T^{-1}(C(\underline{z},t))}(x_n) \right|$$

The sets

$$T^{-1}(C(\underline{z},t)) = \{(\theta,\varphi) \in [0,1]^2 : T(\theta,\varphi) \in C(\underline{z};t)\}$$

are



The boundary curve is piecewise smooth and convex/concave

Def.: The isotropic discrepancy J_N of $P_N = \{\underline{x}_0, \underline{x}_1, \dots, \underline{x}_{N-1}\} \subseteq [0,1]^2$ is defined as

$$J_N(P_N) = \sup_{A \text{ convex}} \left| \lambda(A) - \frac{1}{N} \sum_{n=0}^{N-1} 1_A(\underline{x}_n) \right|.$$

Lemma: We have

$$L_\infty(T(P)) \leq 11\, J_N(P).$$

This is proven by a careful analysis of the shape of the sets $T^{-1}(C(\underline{z};t))$.

<u>Theorem</u>: Let P be a $(0,m,2)$-net in base b. Then

$$J_N(P) \le \frac{4\sqrt{2b}}{\sqrt{N}}$$

<u>Proof idea</u>: Divide the square in subsquares

$$\left[\frac{a}{b^k}, \frac{a+1}{b^k}\right) \times \left[\frac{c}{b^k}, \frac{c+1}{b^k}\right)$$

where $k = \lfloor \frac{m}{2} \rfloor$. The length of the diagonal is $\sqrt{2}\, b^{-k}$

Let $\bar{W}$ be the union of subcubes having nonempty intersection with A or its boundary and W° be the union of subcubes fully contained in A.

We have

$$\frac{1}{N}\sum_{n=0}^{N-1} 1_{\bar{W}}(\underline{x}_n) = \lambda(\bar{W}), \quad \frac{1}{N}\sum_{n=0}^{N-1} 1_{W^\circ}(\underline{x}_n) = \lambda(W^\circ).$$

Then

$$\frac{1}{N}\sum_{n=0}^{N-1} 1_A(\underline{x}_n) - \lambda(A) \le \frac{1}{N}\sum_{n=0}^{N-1} 1_{\bar{W}}(\underline{x}_n) + \lambda(\bar{W}\setminus A) - \lambda(\bar{W}) = \lambda(\bar{W}\setminus A),$$

$$\frac{1}{N}\sum_{n=0}^{N-1} 1_A(\underline{x}_n) - \lambda(A) \ge \frac{1}{N}\sum_{n=0}^{N-1} 1_{W^\circ}(\underline{x}_n) - \lambda(W^\circ) - \lambda(A\setminus W^\circ) = -\lambda(A\setminus W^\circ)$$

Now $\bar{W}\setminus A \subseteq \{\underline{x} \in [0,1]^2 \setminus A : \|\underline{x} - \underline{y}\| \le \sqrt{2}\, b^{-k} \text{ for some } \underline{y} \in A\}$.

Since the set A is convex, the length of its boundary is less then the length of the boundary of the square, which is 4.

Thus $\lambda(\bar{W}\setminus A) \le 4\sqrt{2}\, b^{-k}$.

Similarly

$$A \setminus W^o \subseteq \{ \underline{x} \in A : \| \underline{x} - \underline{y} \| \leq \sqrt{2}\, b^{-k} \text{ for some } \underline{y} \in [0,1]^2 \setminus A \}$$

Thus $\lambda(A \setminus W^o) \leq 4\sqrt{2}\, b^{-k}$.

$\square$

(3) The acceptance-rejection sampler

(3.1) The AR algorithm

Consider the problem of sampling from a distribution whose probability density function is only known partially, i.e. let

$$\psi : [0,1] \to \mathbb{R}_+$$

such that $\int_0^1 \psi(x)\,dx > 0$.

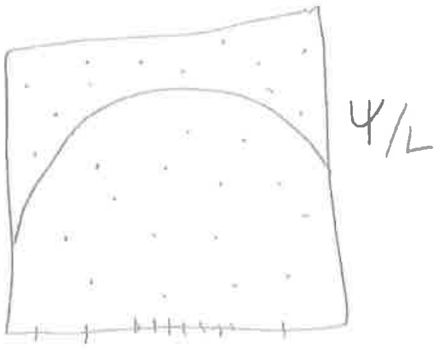
We want to generate low-discrepancy samples with pdf

$$\frac{\psi}{\int_0^1 \psi(x)\,dx}.$$

One way of doing so is the following: Let $L>0$ such that $\psi(x) \le L$ $\forall x \in [0,1]$. Then generate a uniformly distributed point in $[0,1]^2$, say (u_1, u_2). Then if

$$u_2 \le \frac{\psi(u_1)}{L},$$

we accept the point u_1 as a sample, otherwise reject.



We want to use this algorithm to generate low-discrepancy samples $P = \{x_0, x_1, \ldots, x_{N-1}\}$:

$$D(P) = \sup_{0 \le t \le 1} \left| \frac{1}{N} \sum_{n=0}^{N-1} 1_{[0,t]}(x_n) - \frac{\int_0^t \psi(x)\,dx}{\int_0^1 \psi(x)\,dx} \right|$$

That is we generate a $(0, m, 2)$-net $Q = \{\underline{y}_0, \underline{y}_1, \ldots, \underline{y}_{M-1}\}$, $M = b^m$.
Then let

$$P = \left\{ y_{n,1} \in [0,1] : y_{n2} \le \frac{\psi(y_{n,1})}{L}, \ 0 \le n \le b^m - 1 \right\}.$$

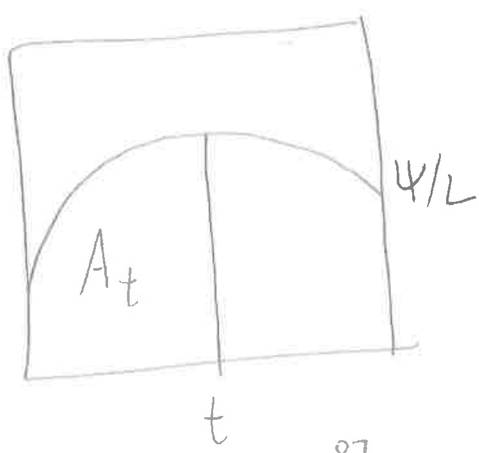
We have

$$D(P) = \sup_{0 \le t \le 1} \left| \frac{1}{N} \sum_{n=0}^{N-1} 1_{[0,t]}(x_n) - \frac{\int_0^t \psi(x)\,dx}{\int_0^1 \psi(x)\,dx} \right|$$

$$= \sup_{0 \le t \le 1} \left| \frac{1}{N} \sum_{n=0}^{M-1} 1_{A_t}(y_n) - \frac{\lambda(A_t)}{\lambda(A_1)} \right|,$$

where

$$A_t = \left\{ (y_1, y_2) \in [0,1]^2 : y_2 \le \frac{\psi(y_1)}{L} \right\}.$$



Now we have

$$\left| \frac{1}{N} \sum_{n=0}^{M-1} 1_{A_t}(y_n) - \frac{\lambda(A_t)}{\lambda(A_1)} \right|$$

$$\leq \frac{M}{N} \left| \frac{1}{M} \sum_{n=0}^{M-1} 1_{A_t}(\underline{y}_n) - \lambda(A_t) \right| + \left| \overbrace{\lambda(A_t)}^{\leq \lambda(A_1)} \left(\frac{M}{N} - \frac{1}{\lambda(A_1)} \right) \right|$$

$$\leq \frac{M}{N} \left[\left| \frac{1}{M} \sum_{n=0}^{M-1} 1_{A_t}(\underline{y}_n) - \lambda(A_t) \right| + \left| \lambda(A_1) - \frac{1}{M} \underbrace{\sum_{n=0}^{M-1} 1_{A_1}(\underline{y}_n)}_{=N} \right| \right]$$

$$\leq 2 \frac{M}{N} \sup_{0 \leq t \leq 1} \left| \frac{1}{M} \sum_{n=0}^{M-1} 1_{A_t}(\underline{y}_n) - \lambda(A_t) \right|$$

$\leq 2 \frac{M}{N} J_M(Q)$, provided that ψ is ~~convex or~~ concave (or convex).

Thus

$$D_N(P) \leq 2 \frac{M}{N} J_M(Q).$$

if ψ is concave (or convex).

Note that as $M \to \infty$, we have $\frac{N}{M} \to \lambda(A_1)$. Thus, for M large enough, we get

$$D_N(P) \leq \frac{4}{\lambda(A_1)} J_M(Q)$$

using $\frac{M}{N} \leq \frac{2}{\lambda(A_1)}$

Thus if we use a (digital) $(0,m,2)$-net we obtain

$$D_N(P) \le \frac{4}{\lambda(A_1)} \frac{4\sqrt{28}}{\sqrt{M}} \le \frac{16\sqrt{56}}{\sqrt{\lambda(A_1)}} \frac{1}{\sqrt{N}},$$

assuming that M is large enough such that

$$\frac{\lambda(A_1)}{2} \le \frac{N}{M} \le 2\lambda(A_1).$$

Theorem: Let Q be a digital $(0,m,2)$-net and let ψ be a concave (unnormalized) densitiy function. Let P be the set accepted by the acceptance-rejection algorithm. Then

$$D(P) \le \frac{c}{\sqrt{N}},$$

where N is the number of accepted points.

This also works in dimension s, i.e. $\psi : [0,1]^{s-1} \to \mathbb{R}_+$, in which case we get a convergence rate of order $N^{-\frac{1}{s}}$.

There is also a lower bound.

Theorem: Let $Q \subseteq [0,1]^s$ be an arbitrary point set. Then there exists a convex density function $\psi : [0,1]^{s-1} \to \mathbb{R}_+$ such that for the point set P generated by the AR algorithm, we have

$$D(P) \ge c_s N^{-\frac{2}{s+1}}.$$

Proof idea as for lower bound on isotropic discrepancy.

For $s=2$, i.e. the interval $[0,1)$, one can get a better rate of convergence, assuming that ψ is twice continously differentiable and $0<a \leq |\psi''(x)| \leq A$ for all $x \in (0,1)$. In this case

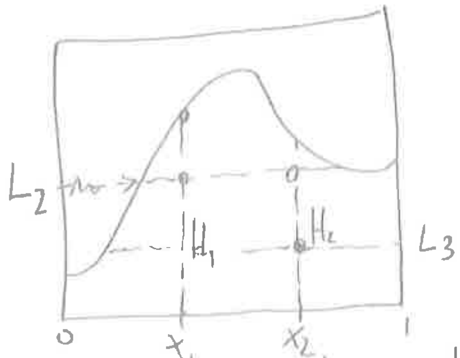
$$D(P) \lesssim N^{-\frac{2}{3}}$$

where Q is a Fibonacci rule. (Houying Zhu talk.)

- Extension to $\mathbb{R}^s$ is possible.
- Remove convexity assumption? Log concave functions? $N^{-\frac{2}{3}}$?
- Problem: acceptance rate can be low, especially in high dimensions;

To avoid the problem of low-acceptance rate one can use the slice sampler.

Let $\psi : [0,1] \to \mathbb{R}_+$ be an unnormalized density.



generate $x_1 \sim U[0,1]$. Generate $y_1 \sim U([0, \psi(x_1)])$. Generate x_2 uniformly on L_2. Generate $y_2 \sim U([0, \psi(x_2)])$. Generate x_3 uniformly on L_3. And so on.

Use the samples $x_1, x_2, x_3, \ldots$ as samples from $\frac{\psi}{\int \psi}$.

Again we would like bounds on the discrepancy of $\{x_1, x_2, x_3, \dots, x_N\} = P$

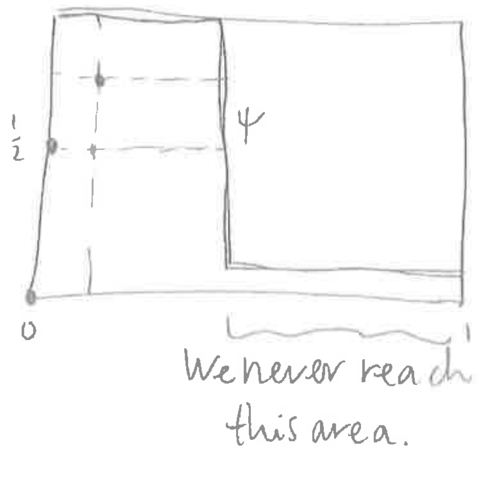
$$D(P) = \sup_{0 \le t \le 1} \left| \frac{1}{N} \sum_{n=0}^{N-1} 1_{[0,t)}(x_n) - \frac{\int_0^t \psi(x)\,dx}{\int_0^1 \psi(x)\,dx} \right| .$$

Note that uniform distribution in $[0,1]$ of the driver sequence $x_1, y_1, x_2, y_2,$ ~~(x_n, y_n)~~, $1 \le n \le M$ is not enough to guarantee convergence.

Let $x_1, y_1, x_2, y_2, \dots$ be the points taken from the van der Corput sequence: $n = n_0 + n_1 b + \dots$, then $\varphi_b(n) = \frac{n_0}{b} + \frac{n_1}{b^2} + \dots$: In $b=2$:

$0, \frac{1}{2}, \frac{1}{4}, \frac{3}{4}, \frac{1}{8}, \frac{5}{8}, \frac{3}{8}, \frac{7}{8}, \dots$

Consider



Slice sampler based on van der Corput sequence does not even converge.

Another problem is the following:

In some cases only the Fourier transform is known.
Say we have a pdf of the form $f: \mathbb{R} \to \mathbb{R}_+$ and we know

$$F(k) = \int_{\mathbb{R}} f(x) e^{-2\pi i k x} \, dx$$

but not f. Hence $f(x) = \int_{\mathbb{R}} F(k) e^{2\pi i k x} \, dk$.

How can we generate low discrepancy samples from f?

The books

- Luc Devroye, Non-Uniform Random Variate Generation; Springer, 1986
- Wolfgang Hörmann, Josef Leydold and Gerhard Derflinger, Automatic Nonuniform Random Variate Generation; Springer, Berlin, 2004

contain an abundance of algorithms where one can try to replace pseudo-random numbers with suitable QMC point sets to ~~achieve~~ obtain samples which have small discrepancy with respect to a target density.

(4) Epilogue

QMC has many applications.

- Machine learning: compute feature maps; Bayesian integration
- Statistics: maximum likelyhood estimation
- SDE and PDE with random coefficients
- Uncertainty Quantification (UQ)

UQ is a huge area which is in need of computational techniques for high dimensional numerical integration.

Important questions for QMC in UQ are:

- Strong tractability; Information based complexity
- Integration over $\mathbb{R}^s$ with respect to normal distr., log-normal distr., including higher order methods and in combination with tractability questions; optimal convergence rates;
- Numerical performance; empirical convergence rate, fast QMC matrix vector product;
- Function approximation rather than integration; tractability and optimal convergence rates;
- Bochner integration and approximation (Hilbertspace valued functions)

(The setting can be abstract, but should be defined in such a way that it can be used in UQ or other application areas.)

- Can we efficiently sample from a pdf p, which is the solution of a PDE, with low discrepancy.
- For the statistics problem and bounds on the spherical cap discrepancy of points lifted from the square to the sphere, it would be of interest to have explicit constructions of points in the square with small discrepancy with test sets as convex sets with smooth boundary and bounded curvature.