

Distribuições Regulares de Pontos sobre Superfícies de Hiperesferas

Lisandro Lovisolo

PROSAICO - DETEL/PEL - UERJ

Agradeço ao professor Eduardo A. B. da Silva (UFRJ) por ter me mostrado, iniciado, e trabalhado comigo neste tema.

September 15, 2023

1 – Problema

- Como gerar conjuntos de K pontos uniformemente (regularmente) distribuídos numa N -esfera?
- Seja $\mathbf{x} = [x_0, \dots, x_{N-1}] \in \mathbb{R}^N$, a N -esfera é o conjunto tal que

$$\ell_2(\mathbf{x}) = \left(\sum_0^{N-1} x_i^2 \right)^2 = r. \quad (1)$$

- Como definir uniformidade do conjunto de pontos $\{\mathbf{x}^{(k)}\}_{k=0 \dots K-1}$?
 - Estocasticamente?
 - Geometricamente?

- O problema de Tammes considera o empacotamento de um determinado número de pontos na superfície de uma esfera de modo que a distância mínima entre os pontos seja maximizada.
- Em geometria e teoria de códigos, um código esférico com parâmetros (N, K, r) é um conjunto de K pontos na hipersfera unitária em N dimensões tal que o produto escalar entre os vetores unitários da origem a quaisquer dois pontos do código é menor ou igual a r .
- O problema do número de beijos (*kissing number*) considera encontrar K máximo, dado N para o qual existe um código esférico com parâmetros (N, K, r) .
- Repara-se que o problema de Tammes consiste em encontrar um código esférico com r mínimo para N e K dados.
 - Pode ser visto como um caso especial particular do problema generalizado de Thomson de minimizar a força Coulomb total entre os elétrons em um arranjo esférico.
 - Há soluções comprovadas apenas para um pequeno número de círculos: 3 a 14 e 24.
 - Há soluções conjecturadas para muitos outros casos, incluindo dimensões superiores.
- Tais conjuntos de pontos possuem diversas aplicações.
- Mas não há solução genérica.

2 – Em 2D, as Soluções são Triviais

2.1 – Abordagem Geométrica

- Divide-se a circunferência em tantos arcos quantos pontos se deseja, obtendo o comprimento de arco

$$\theta_K = \frac{2\pi}{K}.$$

- Arbitra-se um ponto $\mathbf{x}^{(0)}$ sobre a circunferência

$$\mathbf{x}^{(0)} \text{ é tal que } \|\mathbf{x}^{(0)}\| = 1, \theta^{(0)} = \arctan\left(\frac{x_0}{x_1}\right).$$

- Para $k = 1 \dots K - 1$
 - Posiciona-se o ponto $\mathbf{x}^{(k)}$ a θ radianos de $\mathbf{x}^{(k-1)}$, i.e., faz-se

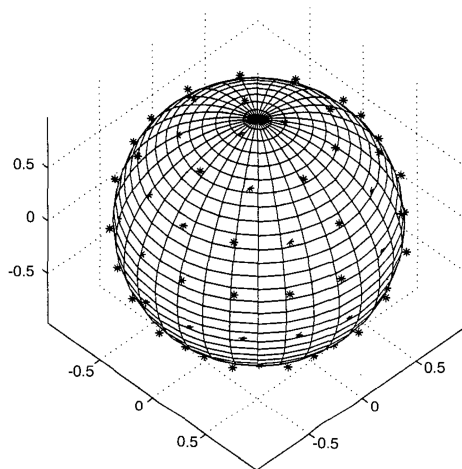
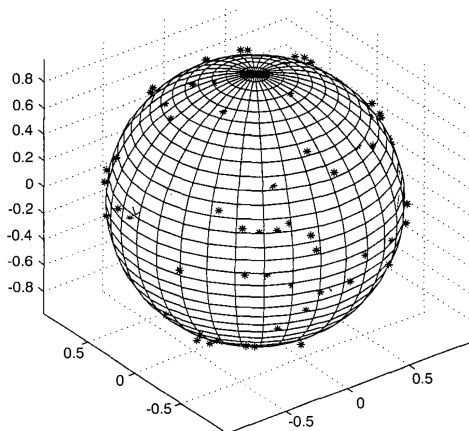
$$\theta^{(k)} = \theta^{(k-1)} + \theta_K, \text{ com } \|\mathbf{x}^{(k)}\| = 1.$$

2.2 – Abordagem Estocástica

- Se x_0 e x_1 são VAs Gaussianas i.i.d. (independentes e identicamente distribuídas), então a VA gerada pelo razão entre elas é uniforme em $[0, 2\pi]$.
 - Assim, sorteando K vetores Gaussianos i.i.d., temos K pontos uniformemente distribuídos na circunferência no sentido estocástico.
 - Mas não há garantia de regularidade nem de minimização da energia (maximização do ângulo ou do arco, distância, entre dois pontos quaisquer).
- ⇒ Claramente as distribuições são distintas e com sentidos distintos!

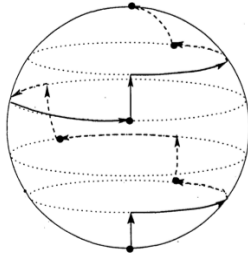


2.3 – Visualizando em 3D

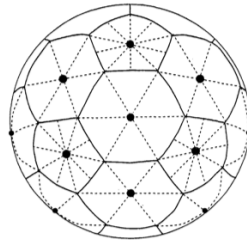


3 – Abordagens Geométricas

- Há algumas abordagens geométricas assintoticamente ótimas para $N = 3$.
- Mas não há método generalizado para $N > 3$.



Alocação de 6 pontos via
espirais generalizadas



Equilíbrio de elétrons
Células de Dirichlet
Bola de *football*

4 – Abordagens Algébricas

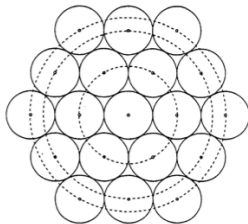
- Para algumas quantidades de pontos K e dimensões N específicas há soluções obtidas a partir de camadas de reticulados regulares.

Packing: Alocação de K pontos sobre a N -esfera de forma a maximizar a distância mínima entre eles (ou equivalentemente o ângulo mínimo entre eles).

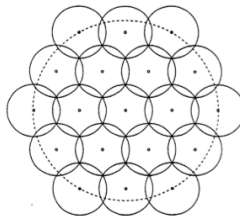
O empacotamento procura por arranjos de esferas sobre uma superfície, de modo que não ocorra nenhuma sobreposição e tal que se qualquer círculo for ampliado ocorrerá sobreposição.

Covering: Alocação de K pontos sobre a N -esfera de forma a minimizar a distância máxima entre um ponto qualquer sobre a N -esfera e ponto mais próximo dentre K pontos alocados.

A cobertura procura por arranjos de esferas sobrepostas que cobrem todo o espaço.



Packing

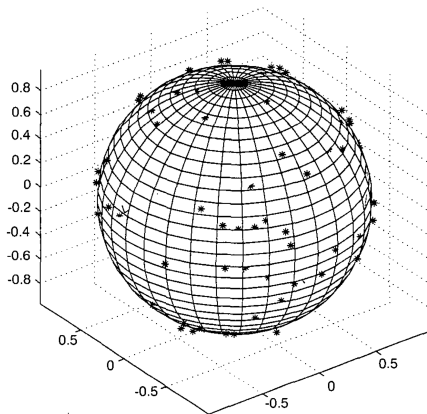


Covering

5 – Abordagens Estocásticas

- Para $N = 3$ e $N = 4$ há métodos usando distribuições uniformes, testes de valores das coordenadas do vetor, seguido de transformações.
- Extensão do método 2D para $N = 3$ e $N = 4$ usando distribuições uniformes.
 - $x_0 \dots x_{N-1}$ são VAs Gaussianas i.i.d.
 - Normaliza-se os vetores usando

$$\frac{1}{\sqrt{x_0^2 + \dots + x_{N-1}^2}} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N-1} \end{bmatrix}$$



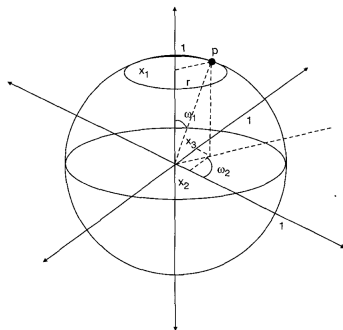
6 – Nosso Método

6.1 – Representação em Coordenadas Polares

- Um ponto p na sobre a N -esfera é dado por

$$p = (1, \omega_1, \omega_2, \dots, \omega_{N-1})$$

$$\text{com } \begin{cases} 0 \leq \omega_n \leq \pi, & n = 1 \dots N - 2 \\ 0 \leq \omega_n \leq 2\pi, & n = N - 1 \end{cases}$$



$$x_1 = \cos \omega_1$$

$$x_2 = \sin \omega_1 \cos \omega_2$$

$$x_3 = \sin \omega_1 \sin \omega_2 \cos \omega_3$$

$$\vdots$$

$$x_{N-2} = \sin \omega_1 \sin \omega_2 \dots \sin \omega_{N-3} \cos \omega_{N-2}$$

$$x_{N-1} = \sin \omega_1 \sin \omega_2 \dots \sin \omega_{N-2} \cos \omega_{N-1}$$

$$x_N = \sin \omega_1 \sin \omega_2 \dots \sin \omega_{N-2} \sin \omega_{N-1}$$

6.2 – Ideia Central

- Quando K é suficientemente grande, temos um revestimento da N -esfera por hipercubos de dimensão $N - 1$.
- Se o comprimento do hipercubo é δ temos as relações

$$\delta = \Delta\omega_1,$$

$$\delta = \Delta\omega_j \prod_{i=1}^{j-1} \sin \omega_i, \quad j = 1 \dots N - 1 \quad [6].$$

- A partir do que, podemos obter os comprimentos dos arcos que geram o hipercubo de dimensões $N - 1$ de lado δ .

- Logo, dado δ e um ponto $(1, \omega_1, \omega_2, \dots, \omega_{N-1})$ qualquer, podemos obter, recursivamente, K pontos via

```
for  $\omega_1 = \Delta\omega_1/2$  to  $\pi$  in increments of  $\Delta\omega_1$  do  
  compute  $\Delta\omega_2$  from eqn. 6  
  for  $\omega_2 = \Delta\omega_2/2$  to  $\pi$  in increments of  $\Delta\omega_2$  do  
    .  
    .  
    .  
    compute  $\Delta\omega_{N-2}$  from eqn. 6  
    for  $\omega_{N-2} = \Delta\omega_{N-2}/2$  to  $\pi$  in increments of  $\Delta\omega_{N-2}$   
    do  
      compute  $\Delta\omega_{N-1}$  from eqn. 6  
      for  $\omega_{N-1} = \Delta\omega_{N-1}/2$  to  $2\pi$  in increments of  
       $\Delta\omega_{N-1}$  do  
        for  $i = 1 \dots N - 1$  do  
          compute  $x_i = \sin \omega_1 \dots \sin \omega_{i-1} \cos \omega_i$   
        end  
        compute  $x_N = \sin \omega_1 \dots \sin \omega_{N-2} \sin \omega_{N-1}$   
      end  
    end  
  end  
end
```

6.3 – Determinando δ

- δ depende de K entre a área da N -esfera.
- A princípio, podemos aproximá-lo via

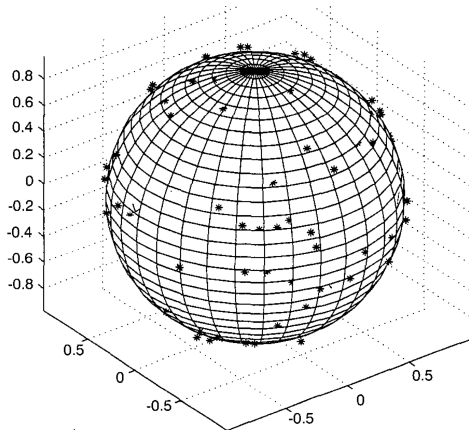
$$A_N = K\delta^{N-1}$$
$$\delta = \left(\frac{A_N}{K}\right)^{\frac{1}{N-1}}.$$

- As suposições de revestimento são somente assintoticamente válidas.
 - Assim, não obtemos o K exato, mas K' ;
 - Iteramos o procedimento fazendo

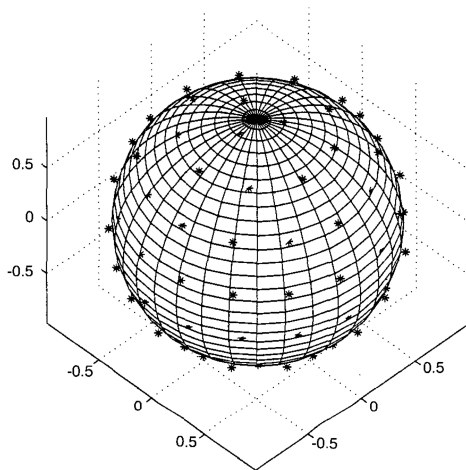
$$\delta \leftarrow \delta \left(\frac{K'}{K}\right)^{\frac{1}{N-1}},$$

até obtermos $K' = K$.

6.4 – Exemplos

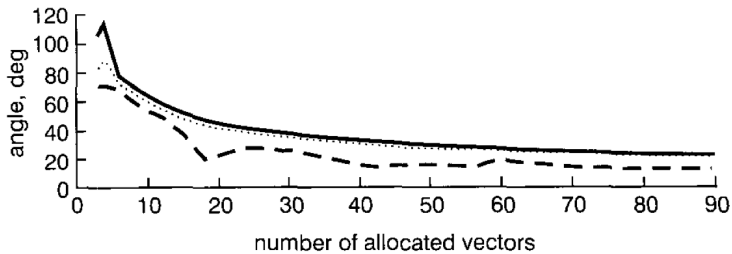


V.A.s IID Gaussianas



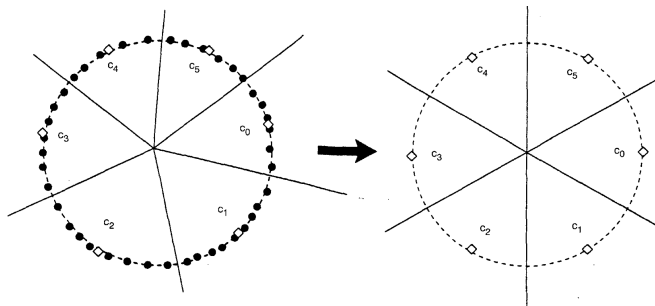
Nosso método

- Comportamento dos ângulos máximo, médio e mínimo entre dois x_k mais próximos em $N = 3$ para $K = 3 \dots 90$.



6.5 – “Treinamento” Estocástico

- Pode-se treinar os K pontos usando $L \gg K$.
- Aplica-se um algoritmo de clusterização/aprendizado: o algoritmo de Linde-Buzo-Gray (LBG).

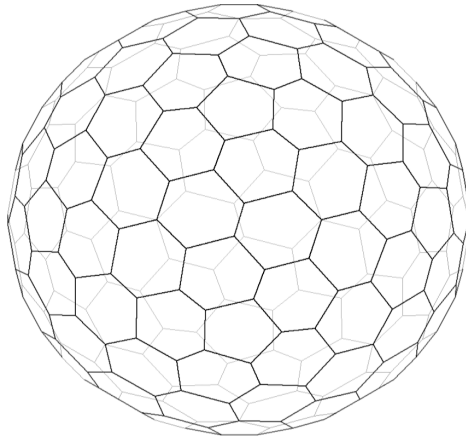


- Ainda assim, L deve ser muito muito grande, e cresce com N .
- Não é efetivo.
- Podemos encontrar alguma distribuição inicial que facilite o trabalho do LBG?

7 – Aplicações

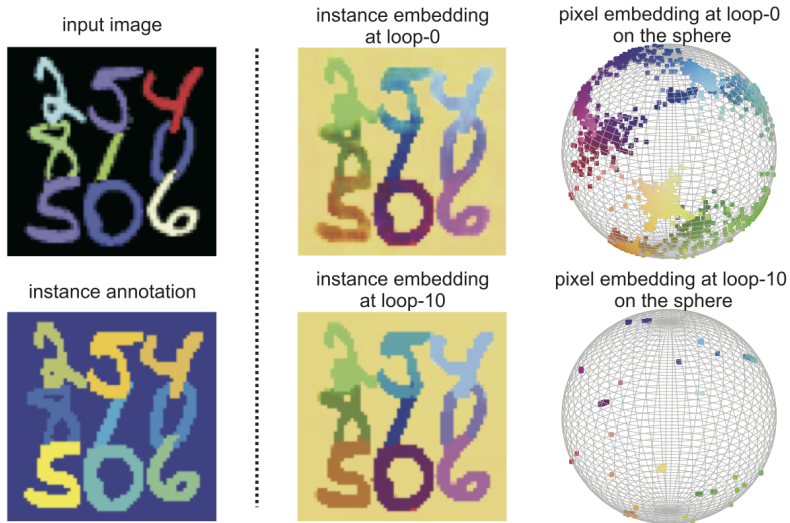
- O problema de encontrar arranjos de K pontos “uniformemente” distribuídos sobre N -esferas tem muitas aplicações além da que me motivou nesse estudo.
- Em simulações computacionais, muitas vezes devem se considerar contornos de modelos e um contorno usando uma N -esfera é isotrópico.
- Algumas aplicações tem usado tais distribuições para obter amostragem esférica:
 - Varredura do interior do modelo a partir de raios oriundos dos pontos desses arranjos.
 - Amostragem para obtenção de estatísticas confiáveis de experimentos confirmatórios.
 - Simulação direcionada com direções determinísticas.
 - Obtenção de *tesellations* (mosaicos).

- ▶ Exemplo de mosaico obtido.



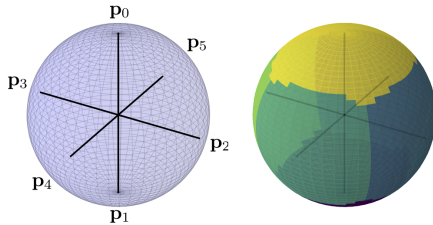
- Algumas aplicações no espaço $SO(3)$ de K orientações discretas distribuídas quase uniformemente na 3-esfera (amostragem esférica):
 - Reconstrução de vírus e culturas de bactérias a partir de experimentos de difração de partículas e raios-X emitidos a partir dos pontos de tais arranjos;
 - Cristalografia;
 - Modelagem de cones do sistema visual humano;
 - Caracterização de câmeras;
- Definição de rotações regularmente uniformes no espaço para:
 - Modelagem da interações e contatos entre objetos sólidos;
 - Modelar as conformações de entidades vivas;
 - Avaliação dos encaixes e interações entre proteínas e fármacos (moléculas).
- Outras aplicações:
 - Livros de códigos para decomposição não-redundante de tensores.
 - Arranjos de pontos de energia mínima.
 - Pontos de partida para otimização por enxame de partículas.
 - Modelagem de receptores com diversidade.

- *Pixel Embeddings* (para separação de classes)

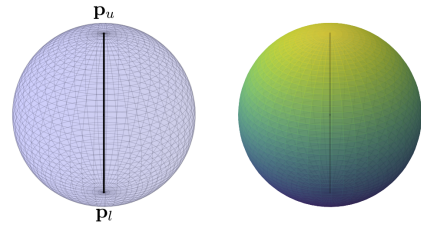


7.0.1– Hyperspherical Prototype Networks

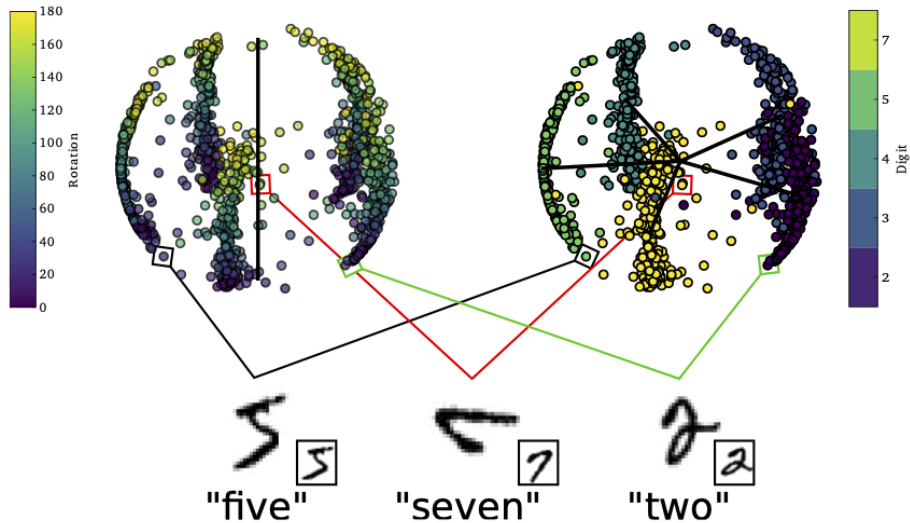
- Advoga-se que usar hiperesferas como os protótipos de espaços de saída para problemas de classificação e regressão.
- Os protótipos de saída não são obtidos a partir de exemplos.



(a) Classification.



(b) Regression.



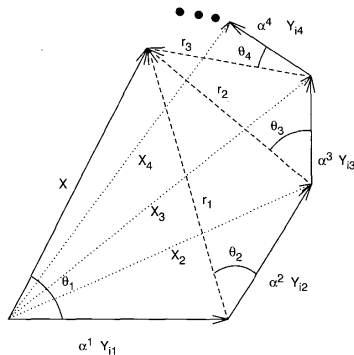
8 – Contexto Histórico (meu)

- Comecei a me interessar por este problema no estudo da codificação/quantização vetorial por aproximação sucessiva.
- Seja vetor \mathbf{x} num espaço N dimensional que devem ser representados por uma soma de vetores predefinidos de um livro-código (*codebook*)

$$\mathbf{C}_{N,K} = \{\mathbf{y}_k\}_{k=1\dots K}, \quad \|\mathbf{y}_k\| = 1, \quad \forall k$$

- Aproximamos \mathbf{x} através de

$$\mathbf{x}_I = \sum_{i=1}^I \alpha^i \mathbf{y}_{k_i}, \quad 0 \leq \alpha \leq 1$$



- Tal representação só é útil se

$$\lim_{I \rightarrow \infty} \| \mathbf{x} - \mathbf{x}_I \| = 0$$

- Seja

$$\Theta(\mathbf{C}_N) = \max_{\|\mathbf{x}\|=1} \left\{ \min_{\mathbf{y}_k \in \mathbf{C}_N} \arccos(\mathbf{x} \cdot \mathbf{y}_k) \right\},$$

então, a quantização vetorial por aproximações sucessivas converge se

$$\alpha \begin{cases} \geq \frac{1}{2 \cos \Theta(\mathbf{C}_N)} & \text{se } \Theta(\mathbf{C}_N) \leq \frac{\pi}{4} \\ \geq 2 \sin \Theta(\mathbf{C}_N) & \text{se } \Theta(\mathbf{C}_N) \geq \frac{\pi}{4} \end{cases}$$

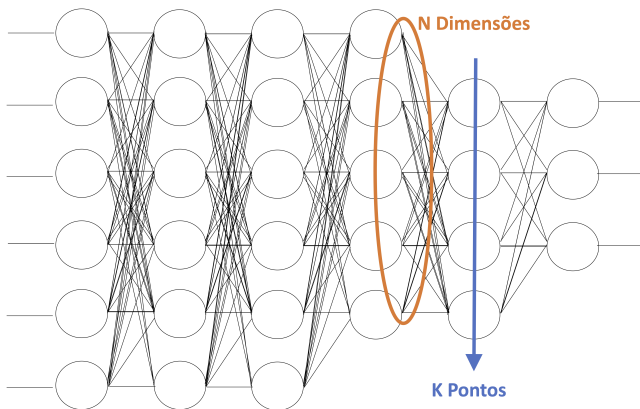
e

$$\| \mathbf{x} - \mathbf{x}_I \| \leq B\alpha^I.$$

- Quanto menor $\Theta(\mathbf{C}_N)$, menor o α que garante a convergência e menor o erro após I passos aproximativos.
- Para reduzir $\Theta(\mathbf{C}_N)$ podemos:
 - Aumentar K ;
 - Distribuir os K vetores o mais uniformemente possível na N -esfera.
 - ▶ No sentido de que os arcos que os ligam tenha o maior comprimento possível!
 - ▶ Isto é, que os ângulos entre eles sejam maximizados.
 - O conceito de uniformidade não está bem definido, obviamente, e este problema equivale à melhor cobertura da N -esfera por $N - 1$ -esferas.

9 – O que Pretendemos Fazer a Seguir?

- Favorecer o processo de treinamento usando pontos mais uniformemente distribuídos capazes de prover embeddings mais separáveis.



- Na camada l temos K_l neurônios, que têm tantas entradas quanto neurônios K_{l-1} existem na camada anterior ($l - 1$).
- Gerar os pesos dos neurônios da camada l a partir de uma distribuição regular de K_l pontos sobre a K_{l-1} -esfera.

Muito obrigado!

lovisolo@eng.uerj.br

10 – Bibliografia

- Problema base:
 - Hicks, J. S., and R. F. Wheeling. "An efficient method for generating uniformly distributed points on the surface of an n-dimensional sphere." *Communications of the ACM* 2.4 (1959): 17-19.
 - Muller, Mervin E. "A note on a method for generating points uniformly on n-dimensional spheres." *Communications of the ACM* 2.4 (1959): 19-20.
 - Marsaglia, George. "Choosing a point from the surface of a sphere." *The Annals of Mathematical Statistics* 43.2 (1972): 645-646.
 - Sloane, Neil JA. "The packing of spheres." *Scientific American* 250.1 (1984): 116-125.
 - Saff, Edward B., and Amo BJ Kuijlaars. "Distributing many points on a sphere." *The mathematical intelligencer* 19 (1997): 5-11.
 - Kunc, Oliver, and Felix Fritzen. "Generation of energy-minimizing point sets on spheres and their application in mesh-free interpolation and differentiation." *Advances in Computational Mathematics* 45.5-6 (2019): 3021-3056.
 - Conway, John Horton, and Neil James Alexander Sloane. *Sphere packings, lattices and groups*. Vol. 290. Springer Science & Business Media, 2013.
- Nossa proposta:
 - Lovisoló, L., and E. A. B. Da Silva. "Uniform distribution of points on a hyper-sphere with applications to vector bit-plane encoding." *IEE Proceedings-Vision, Image and Signal Processing* 148.3 (2001): 187-193.
- Aplicações:
 - Craizer, Marcos, Eduardo Antônio Barros da Silva, and E. G. Ramas. "Convergent algorithms for successive approximation vector quantisation with applications to wavelet image compression." *IEE Proceedings-Vision, Image and Signal Processing* 146.3 (1999): 159-164.

- Sato, Katsushi, and Atsushi Yamaji. "Uniform distribution of points on a hypersphere for improving the resolution of stress tensor inversion." *Journal of structural geology* 28.6 (2006): 972-979.
- Solernou, Albert, and Juan Fernandez-Recio. "Protein docking by Rotation-Based Uniform Sampling (RotBUS) with fast computing of intermolecular contact distance and residue desolvation." *BMC bioinformatics* 11 (2010): 1-14.
- Saldin, D. K., et al. "Reconstructing an icosahedral virus from single-particle diffraction experiments." *Optics express* 19.18 (2011): 17318-17335.
- Moths, Brian, and Abbas Ourmazd. "Bayesian algorithms for recovering structure from single-particle diffraction snapshots of unknown orientation: a comparison." *Acta Crystallographica Section A: Foundations of Crystallography* 67.5 (2011): 481-486.
- Finlayson, Graham D., et al. "Spectral sharpening by spherical sampling." *JOSA A* 29.7 (2012): 1199-1210.
- Papaioannou, Iason. Non-intrusive finite element reliability analysis methods. Diss. Technische Universität München, 2012.
- Giannakis, Dimitrios, Peter Schwander, and Abbas Ourmazd. "The symmetries of image formation by scattering. I. Theoretical framework." *Optics express* 20.12 (2012): 12799-12826.
- Schwander, Peter, et al. "The symmetries of image formation by scattering. II. Applications." *Optics express* 20.12 (2012): 12827-12849.
- Dashti, Ali. Extracting the structure and conformations of biological entities from large datasets. Diss. The University of Wisconsin-Milwaukee, 2013.
- Kyrgyzov, Olexiy, and Deniz Erdogmus. "Nonnegative non-redundant tensor decomposition." *Frontiers of Mathematics in China* 8 (2013): 41-61.
- Vazquez-Corral, Javier, David Connah, and Marcelo Bertalmío. "Perceptual color characterization of cameras." *Sensors* 14.12 (2014): 23205-23229.

- Donatelli, Jeffrey J., and James A. Sethian. "Algorithmic framework for X-ray nanocrystallographic reconstruction in the presence of the indexing ambiguity." *Proceedings of the National Academy of Sciences* 111.2 (2014): 593-598.
- Nuwanpriya, Asanka, Siu-Wai Ho, and Chung Shue Chen. "Indoor MIMO visible light communications: Novel angle diversity receivers for mobile users." *IEEE Journal on selected areas in communications* 33.9 (2015): 1780-1792.
- Zhu, Bingcheng, et al. "Performance bounds for diversity receptions over arbitrarily correlated Nakagami- m fading channels." *IEEE Transactions on Wireless Communications* 15.1 (2015): 699-713.
- Pande, K., et al. "Simulations on time-resolved structure determination of uncrystallized biomolecules in the presence of shot noise." *Structural Dynamics* 2.2 (2015).
- Kong, Shu, and Charless C. Fowlkes. "Recurrent pixel embedding for instance grouping." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.
- Yu, Wenbo, et al. "Exploring protein-protein interactions using the site-identification by ligand competitive saturation methodology." *Proteins: Structure, Function, and Bioinformatics* 87.4 (2019): 289-301.
- Stevens, Nathaniel T., and Christine M. Anderson-Cook. "Design and analysis of confirmation experiments." *Journal of Quality Technology* 51.2 (2019): 109-124.
- Kunc, Oliver, and Felix Fritzen. "Generation of energy-minimizing point sets on spheres and their application in mesh-free interpolation and differentiation." *Advances in Computational Mathematics* 45.5-6 (2019): 3021-3056.
- Phutke, Minakshee, Jenil Dedhia, and A. K. Suresh. "Modelling solid-solid reactions: Contact-point approach." *Chemical Engineering Journal* 377 (2019): 120570.
- Mettes, Pascal, Elise Van der Pol, and Cees Snoek. "Hyperspherical prototype networks." *Advances in neural information processing systems* 32 (2019).
- Jo, Sunhwan, et al. "Computational Characterization of Antibody-Excipient Interactions for Rational Excipient Selection Using the Site Identification by Ligand Competitive Saturation-Biologics Approach." *Molecular pharmaceutics* 17.11 (2020): 4323-4333.

- Wang, Zhennan, et al. "MMA regularization: Decorrelating weights of neural networks by maximizing the minimal angles." *Advances in Neural Information Processing Systems* 33 (2020): 19099-19110.
- Motta Goulart, Renan, Carlos Cristiano Hasenclever Borges, and Raul Fonseca Neto. "An evolutionary algorithm for large margin classification." *Soft Computing* 25.11 (2021): 7593-7607.
- Somani, Sandeep, et al. "Toward biotherapeutics formulation composition Engineering using site-identification by ligand competitive saturation (SILCS)." *Journal of pharmaceutical sciences* 110.3 (2021): 1103-1110.
- Ragot, Stéphane, and Adriana Vasilache. "Spherical Vector Quantization for Spatial Direction Coding." *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2023.
- Fakorede, Olukorede, et al. "Improving Adversarial Robustness with Hypersphere Embedding and Angular-Based Regularizations." *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2023.