

# Lab 4 Discussion

## PART A: CANN

### CANN:

1. 写报告时先解释1d-CANN的结构，输入是高斯的，突触连接是高斯型的
2. 我们尝试把CANN的距离增大，这样可以同时两个位置输入刺激。控制刺激距离，（与2a）的大小之差，可以看看两个兴奋的变化，是向中间聚拢（距离较小），还是此消彼长，因为CANN有全局抑制功能。
3. 关于参数k的研究：写报告时可以绘制参数k对刺激时的尖峰电压，撤去刺激时的电压的曲线，或许可以尝试观察两个尖峰电压的比例关系。
4. 再次施加刺激时网络状态的平滑转变（相同和不同位置）

### CANN-SAF:

1. 关注m的大小和速度，速度有一定的范围，参考弭元元老师的论文
2. 同时m越大预测的能力越大，但是电压会减小，说明对预测的置信度下降。
3. 如果能找到兴奋的“重心”（因为往往不是高斯的），得到和刺激的中心差距，绘制关于m的变化曲线。
4. 思考v可以怎样讨论  
(问助教：为什么没有课堂上的左右震荡的现象，预测之后被“拉回来”)。

## PART B: Discussion

### 论文解读：

### Essay: Learning a Continuous Attractor Neural Network from Real Images

the network dynamics enables the neural system to reach to the same stationary state

(attractor), once an external input falls into the basin of attraction.

CANNs have been successfully applied to describe the encoding of a number of continuous features in neural systems, such as orientation, moving direction, head direction, and spatial location of objects.

The key property of a CANN is that it holds a continuous family of stationary states, which form an approximately flat sub-manifold of low energy in the network state space; whereas, in the Hopfield network, attractor states are isolated with each other with high-energy barriers.

由大量自然图像预先训练的深度学习生成对象表示（读出层之前的神经活动），这些表示捕获了高级视觉皮层中神经反应的一些统计特征。

Q:为什么正交的Hebb Learning 效果更好？

## 1. Hebb 学习只能学习独立的模式

原因：

### 1. 权重更新机制

Hebb 学习的权重更新规则是：

$$\Delta w_{ij} = \eta \cdot x_i \cdot y_j$$

这种机制依赖于神经元之间的同步活动。如果新输入模式与已有模式相关（线性相关或部分相关），则权重的更新会受到已有模式的干扰。

### 2. 模式干扰 (Interference)

- 如果新模式和已存储模式相关，其在网络权重上会引入冗余表示。
- 冗余会导致网络存储的模式之间产生干扰，从而削弱网络对新模式的学习能力。

### 3. 存储容量限制

Hebb 学习倾向于分配有限的网络权重空间给所有模式。当模式之间相关性较强时，会占用更多的存储资源，进一步降低网络的存储效率。

## 2. 正交 Hebb 学习如何解决相关模式的问题

**核心思想：**

正交 Hebb 学习通过将新模式正交化，确保它与已有模式不相关，从而避免模式之间的干扰。

正交化公式如下：

$$\eta^{p+1} = \xi^{p+1} - \sum_{\mu=1}^p \hat{\eta}^{\mu} (\hat{\eta}^{\mu} \cdot \xi^{p+1}),$$

其中：

- $\eta^{p+1}$ ：新模式的正交化表示。
- $\xi^{p+1}$ ：新输入模式。
- $\hat{\eta}^{\mu}$ ：已有模式的单位向量。
- $(\hat{\eta}^{\mu} \cdot \xi^{p+1})$ ：新模式在已有模式上的投影。

**正交化带来的好处：**

### 1. 去相关性 (Decorrelation)

- 正交化消除了新模式和已有模式之间的相关性。
- 存储的每个模式都彼此独立，避免了模式干扰。

### 2. 解决相关模式的存储问题

- 正交 Hebb 学习允许存储相关模式，因为它能够有效分离相关模式的独立信息。
- 即使新模式与已有模式高度相关，正交化仍能提取新模式中独特的信息。

### 3. 提高存储容量

- 正交化显式增加了模式间的独立性，网络可以有效利用更多的权重空间存储模式。

## 2. 正交 Hebb 学习中的权重直接赋值机制

特点:

- **权重的计算公式:** 正交 Hebb 学习直接通过以下公式计算权重矩阵  $W$ :

$$W_{ij} = \sum_{\mu=1}^P (\hat{\eta}_i^{\mu} \hat{\eta}_j^{\mu} - \delta_{ij} \hat{\eta}_i^{\mu} \hat{\eta}_i^{\mu}),$$

其中:

- $\hat{\eta}^{\mu}$  是模式  $\mu$  的归一化表示。
- $\delta_{ij}$  是 Kronecker delta, 避免对角线上自连接的影响。
- **一次性计算完成:** 所有模式  $\eta^{\mu}$  在训练阶段已经被正交化, 权重矩阵  $W$  的计算是一种离线处理, 无需在线逐步更新。
- **去干扰性:** 正交化消除了模式之间的相关性, 确保每个模式的权重更新独立, 这种权重直接赋值的方式天然避免了传统 Hebb 学习中的干扰问题。

优缺点:

- 优点: 权重计算一次完成, 存储的模式互不干扰, 适合批量存储的离线任务。
- 缺点: 无法动态适应新模式的在线学习需求。

**实现突触传递:**

Denote  $\mathbf{S} = \{S_i\}$ , for  $i = 1, 2, \dots, N$ , to be the network state, with  $S_i$  taking values of  $\pm 1$ . The network dynamics is given by

$$S_i(t+1) = \text{sign}\left(\sum_j W_{ij} S_j\right), \quad (3)$$

and the energy function of the network state is calculated to be

$$E = -\frac{1}{2} \mathbf{S}^T \mathbf{W} \mathbf{S}. \quad (4)$$

在正交 Hebb 学习中, 能量衡量了网络状态对已存储模式的匹配程度, 低能量对应高匹配。

注意这里的S的更新是

**我们的改进:**

之前的firing rate 限制只能有一个波包 (我没太懂, 或许可以从数学上分析一下)

To be biologically more plausible, we extend the above discrete model to be a continuous one, which is written as,

$$\tau \frac{dV_i}{dt} = -V_i + \sum_j W_{ij} g(V_j) + I_i, \quad (6)$$

(希望我们在某一个Pattern的输入后波包的形态不是太简单，不单单有位置和大小信息)

所以我们在生物学中用神经元的电压来代表状态：

$$E = -\frac{1}{2} V^T W V$$

$$E_t = -\frac{1}{2} V_t^T W V_t$$

Q:为什么椅子在开始和结束能量非常高？按说也是两个状态连续的Pattern。

(可以问助教？)

## 实现步骤

### Phase 1:

首先探讨普通的Hebb Learning 结合其实际意义，在SLT和SLD的体现，在相同位置输入和改变位置输入时的平衡时的电压。

1. 在原有的W（距离Gauss）上加上Hebb学习公式，更新W（应该在撤去输入时更新？）
2. 刺激的位置
  - 相同位置两次输入
  - 另一个位置输入
3. 观测指标
  - 稳定时的电压大小

## Phase 2:

Assumption: 更高级的脑区和视皮层 (哪个V?)

1. 实现 Pattern 的schmidt正交化 (首先使用简单Pattern, 有相关性, 不能使用传统的Hebb Rule)
2. 计算和设置 W 矩阵
3. 完成网络构建, 神经元的数量和Pattern的维数相同, 依然利用 BrainPy
4. 完成V的更新的函数
5. 设置输入 (旧Pattern或者新的Pattern)
6. 观测指标
  - 研究能量在时间序列中的变化, 以揭示网络状态的稳定性和学习的有效性。
  - 不同Pattern输入稳定时的位置。

## 附加的讨论:

1. 新增一个模式 (更新W), 观察兴奋所处的位置及其相邻的Pattern的关系 (但是输入已经有了相关性, 好像一定相关)
2. 引入已有的深度神经网络计算现实图片的Pattern (可能实际做出来效果会不好)

## Q&A实现时遇到的问题: