INTRODUCTION

High dimensional statistic → assumption about simplicity of the data → associated algorithms

Theoretical neuroscience → machine learning: Recognition, to which we knew all of the ingredients

And the methods we use to understand these (and to be invented) paved the way of understanding the brain.

OUTLINE: 1. Older but goodies

2. High dimensional statistics: theory and experiments

3. Deep learning: theory and practice

Ref: Statistical Mechanics of

LECTURE I: OLDIES BUT GOODIES

Hodgkin Huxley → quantitative description of action potential.

Hopfield → an idea of memory

Hodgkin Huxley:

DENDRITICS & Soma

+ + + + +

V = 0

Sodium CHANNEL

Out: Na+ Na+ Na+ if for any reason V(a) >

The channel opens

Na+ rushes in

FAST POSITIVE FEEDBACK

Astrocytes CHANNEL

Energy cost to being excitable to maintain the new equilibrium steady state

Na+ + K+ pumping!

SLOW NEGATIVE FEEDBACK

→ we have the minimal ingredients of an excitable system in biology:

\[
\begin{align*}
&\text{Membrane voltage } V(t) \\
&\text{Na	extsuperscript{+} in } \\
&\text{K	extsuperscript{+} out}
\end{align*}
\]
The spike happening somewhere on the membrane will then propagate, through a wave along the membrane.

Increase of $V$ happens at synapse:

\[ \text{Ca}^{2+} \text{ influx in,} \]
\[ \text{vesicles with neurotransmitter open and release} \]
\[ \text{neurotransmitter diffuses} \]
\[ \text{post synaptic neuron open neurotransmitter,} \]
\[ \text{gated ion channels.} \]

So complicated transmission, that actually contains many degrees of freedom that allow a very sensitive fine-tuning of the spikes → store memories for a lifetime?

Then to study the brain, we need to cause neurons... REF: Heidi Hercul, From Hay and Kepner, 1990.

$+\text{time scale separation: spikes } \sim 2 \text{ms.}$

$\text{input } \sim \text{LED mV.}$

1. INTEGRATE AND FIRE: "leaky integration of membrane voltage."

\[ \frac{dV}{dt} = -V + I(t) \]

if $V > \text{Threshold}$ emit a spike, then go back to $V_{thref}$.

- input current $I(t)$ (viewed.)
- input $I(t)$ $\Rightarrow$ Threshold
- $\text{collective modeling}: I(t) = \sum W_j \sum K(t-t_j)$

$\text{SYNAPTIC STRENGTH}$ $\text{time of spike of } j$

Ref: can formulate mean field theory of integrate and fire networks → a bit complicated.

- SPREAD RATE: one more step of conserving.
  \[ r = \text{rate} = \text{spikes per unit time}. \]

Ex: Compute analytically

\[ \frac{dV}{dt} = -V + \sum W_j V_j(t) + I_{\text{ext}} \]

Ref: Front neuron 200 Hz

SPONTANEOUS RANDOM NEURON 31 Hz
**Spin Neuron**

\[ s_i(t+1) = \text{sgn} \left( \sum_j J_{ij} s_j(t) + h_i \right) \]

- Zero-temperature Glauber dyn of spin system.

**Re:** Other separations of time scales with the updating, to which we have on much longer time scales plasticity = changes on the \( J_{ij} / W_{ij} \).

**Hopfield Model**

What is a memory? pattern completion: associate name with picture, stimulus with action.

\[ \text{memories:} \quad \{ s_i \}^N_{i=1} \quad \{ u(i) \}^P_{i=1} \quad \text{we wish that if initializing in } \{ s_i \}^N_{i=1} \text{ system will converge to } \{ s_i \}^N_{i=1} \text{ closing nip the corrupted memory.} \]

- Image of basins of attraction, how to implement that?

G Hopfield realized that one needs a dynamics with decreasing energy to rule out the oscillations \( \rightarrow \) OK Glauber.

\[ E = - \sum_{i,j} J_{ij} s_i s_j \]

\( \rightarrow \) idea that neurons that fire together, wire together (Hebb):

\[ \Delta J_{ij} = s_i s_j \text{ each time see a pattern} \]

G after seeing all the patterns:

\[ J_{ij} = \frac{1}{N} \sum_{\alpha=1}^P s_i^\alpha s_j^\alpha \]

\[ E = - \frac{1}{N} \sum_{\alpha=1}^P \left( \sum_{i,j} s_i^\alpha s_j^\alpha \right) s_i s_j \]

\[ m^\alpha = \frac{1}{N} \sum_i s_i^\alpha \]

magnetization of \( \alpha \)-th memory = overlap

Looking good.

What about influence between memories? so \( m = (m_1, \ldots, m_P) = (1, 0, 0, \ldots) \) stable?

\[ \text{replica calculation:} \]

\[ \frac{0.05}{0.14} \]

\( \text{lowest free energy, also superposition of memories,} \)

\( \text{yet not too problematic, given an initialization close to a pattern} \)

\( \text{Re: Correlated patterns? Frenkel inverse rule?} \)

Is \( s_i^\alpha \) a fixed point of T=0 dynamics? Sufficient constraint to be in retrieval phase!
first pattern: \[ \frac{s_i^2}{\hat{s}_i^2} = \text{sign} \left( \frac{\sum_{j=0}^{N} \sum_{j'} \hat{s}_i s_{j'} \hat{s}_j}{N} \right) \]

\[ = \text{sign} \left( \frac{\sum_{j} \hat{s}_i \hat{s}_j + \sum_{j'} \hat{s}_i \hat{s}_{j'}}{N} \right) \]

\[ = \frac{s_i^2}{\hat{s}_i^2} = \text{sign} \left( \frac{s_i^2 + \eta_i}{\hat{s}_i^2} \right) \text{ with } \eta_i \sim \mathcal{N}(0, \frac{\sigma^2}{N}) \]

About Learning (Unsupervised / Supervised)

- Statistical Structure of Natural Images → Network Connectivity → Network Activity
- What are the rules of these feedback?
- \( \eta^2 \) fixed points?

Different learning tasks:
- Unsupervised learning
- Supervised learning \( x \to y^* \)
- Reinforcement learning (not told \( y^* \)) but tell you whether good/bad direction
- More biologically plausible

UNSUPervised LEARNING: What is Hebbian learning doing?

Imagine streaming inputs \( x = 1 : M \) and having weights changing according to Hebb's rule:

\[ \Delta w_i = \lambda \sum_{x=1}^{M} x_i w(x) \]

Smaller lambda (slow plasticity)

Remember: inputs \( i \) after \( n \) inputs: \( \tilde{w}(n) = \tilde{w}(o) + \sum_{x=1}^{M} \lambda \tilde{w}(x) \rightarrow \frac{\partial \tilde{w}}{\partial t} = \left( \tilde{w}(x) \right)_{P(w)} \]

Plug in \( \tilde{w}(n) = \tilde{w}(o) + \sum_{x=1}^{M} \lambda \tilde{w}(x) \rightarrow \frac{\partial \tilde{w}}{\partial t} = \left( \sum_{x=1}^{M} \lambda \tilde{w}(x) \right)_{P(w)} = \left( \sum_{x=1}^{M} \lambda \tilde{w}(x) \right)_{P(w)} = \left( \sum_{x=1}^{M} \lambda \tilde{w}(x) \right)_{P(w)} \]

Learning driven by second-order correlations for linear neurons