Synaptic Learning Rules

Section 10

Binxu Wang

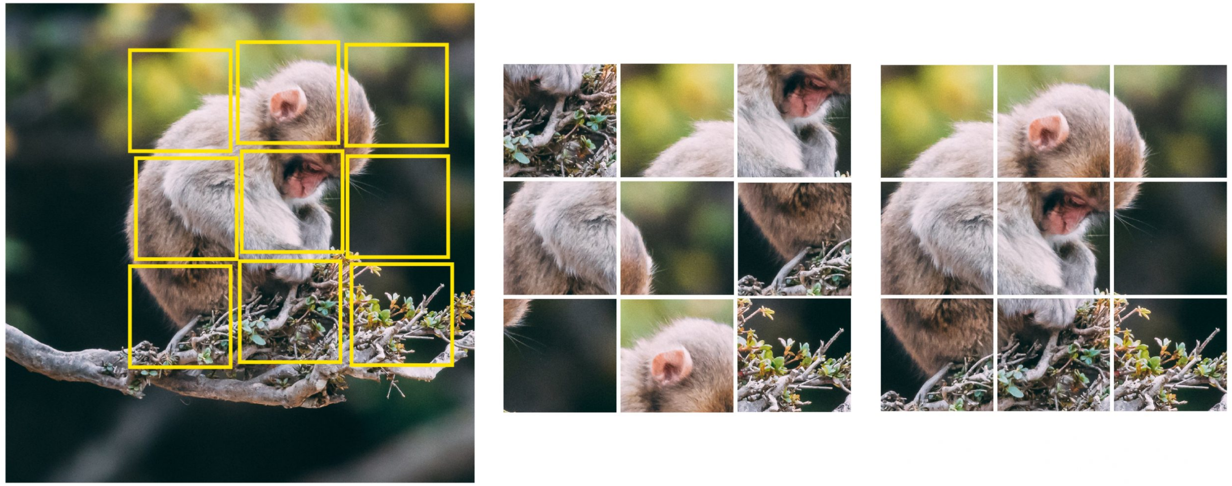
Brief Conceptual Review

Types of Learning in ML

- Supervised
  - You get a set of $\{x_i, y_i\}$
  - Learn a mapping $\hat{y} = f(x)$
  - E.g. regression, LNP model,
- Unsupervised
  - You get a set of $\{x_i\}$
  - Try to learn its distribution $p(x)$.
  - Or to learn **a useful representation** $z = f(x)$.
  - E.g. Dimension reduction
- Reinforcement learning
  - Get an environment.
  - Find out the best way to act in it. Maximize reward.
  - E.g. AlphaGO, AlphaStar

This is classic and general division.
No clear boundary …

- Unsupervised → Self-supervised learning
- Reinforcement → Imitation / Supervised learning
- Supervised learning → RL to maximize some reward for correct prediction

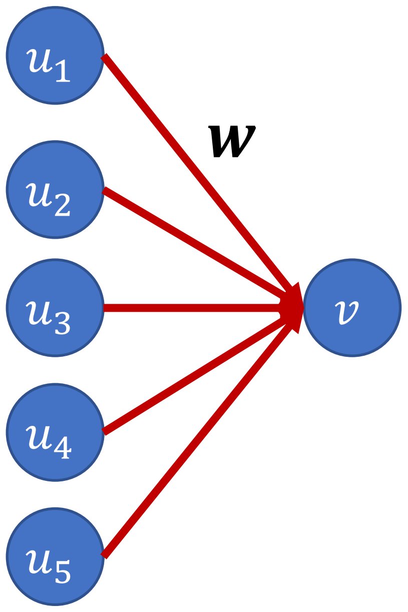


Synaptic Learning Rules

Learning Rules

- Linear neuronal model
$$v = w^T u$$
  - u: pre-synaptic activity
  - v: post-synaptic activity

- Learning rule: How synapses change their strength based on pre-, post- activity?
  - Δw or dw/dt as function of u, v time course
$$\frac{dw}{dt} = \mathcal{F}(u, v)$$



Why *learning rules* are hard to study experimentally?

Timescale of learning

- Months or years.

Phase of learning

- Early in life.

Spatial scale (synapse, neuronal processes)

- Accessible to electron microscopy and other high precision imaging.
- Usually need to sacrifice animal first.

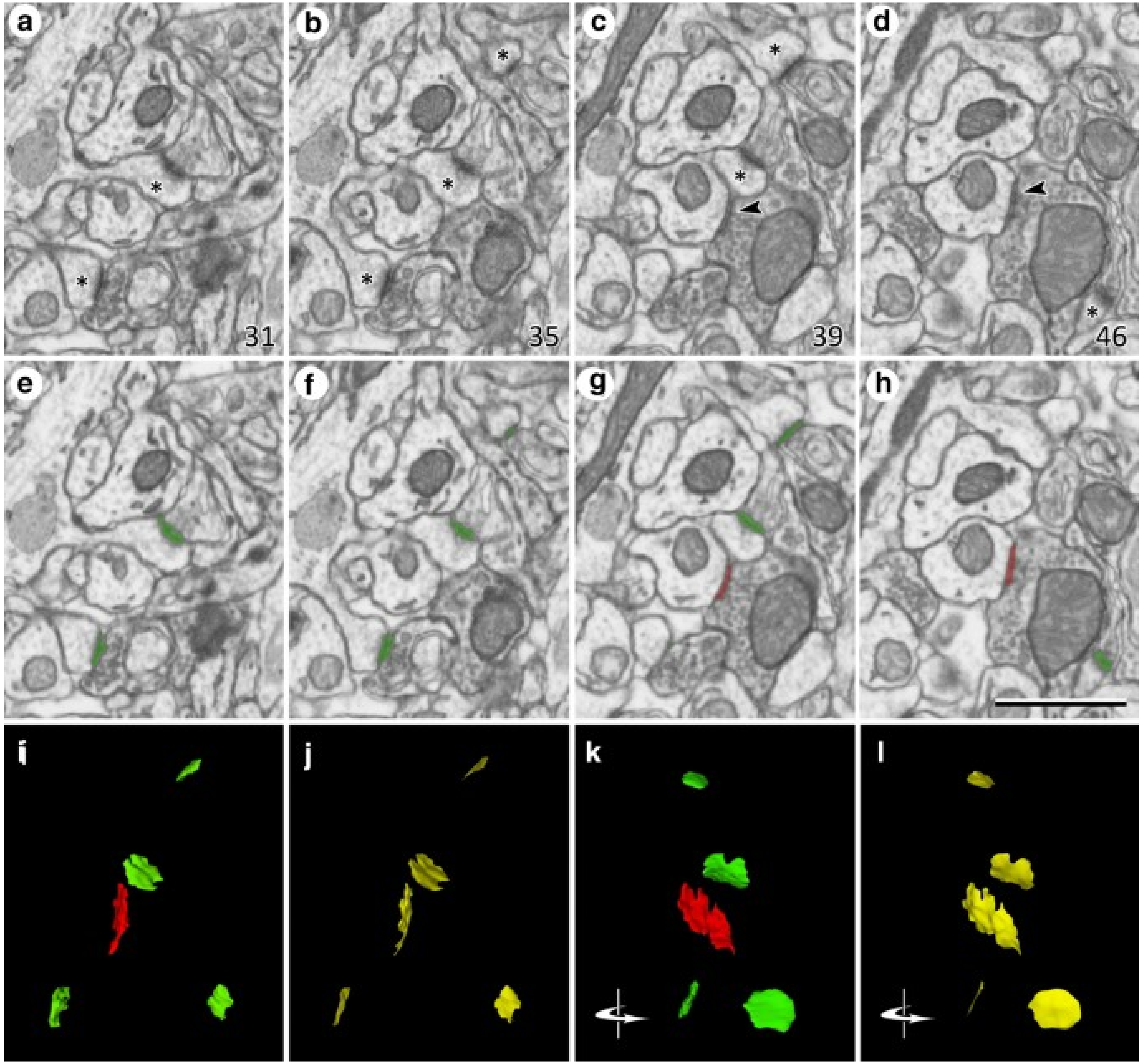
Different rules can come to same behavior/weights

Hard to analyze data or associate it to behavior.

Better theories are required to produce hypothesis to test.

Synaptic Weight as Seen through Electron Microscopy

- Synapses are much smaller than neurons…

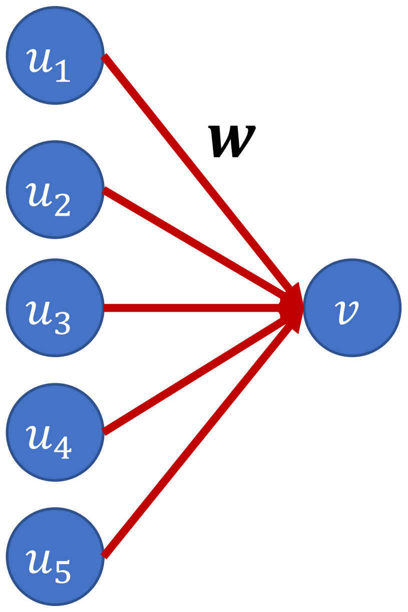


Hebbian Learning

- Simple notion: *Fire together, wire together.*

$$\tau \frac{dw}{dt} = uv$$

- Intuition:
  - All synapse gets stronger. (*LTP*)
  - Pre-synaptic neuron that fires more grows stronger synapse.



Weight Dynamics as Linear Dynamic System

$$\tau \frac{dw}{dt} = (w^T u)u = (uu^T)w$$

- We transform the equation to a linear dynamic system.

$$\frac{dx}{dt} = Ax$$

- Stability criterion for a linear dynamic system?
  - Real part of A eigenvalues are less than (or equal to) 0.

What are the eigenvectors and eigenvalues of uu^T

- $\|u\|^2, 0.$
- u, all others.

Original Hebbian learning rule is not stable!

Weight Norm Analysis

- Hebbian Learning rule

$$\tau \frac{dw}{dt} = uv$$

- Since $u = w^T v$

$$\tau \frac{dw}{dt} = (w^T u)u = (uu^T)w$$

- Dynamics of weight norm

$$\tau w^T \frac{dw}{dt} = w^T (uu^T) w = (u^T w)^2$$

$$\frac{1}{2} \tau \frac{d\|w\|^2}{dt} = (u^T w)^2 \geq 0$$

  - Either no learning or explode in long term.

This general philosophy is to reduce a high dim dynamic system into a low dimensional one, easier to analyze.

$$w \rightarrow \|w\|$$

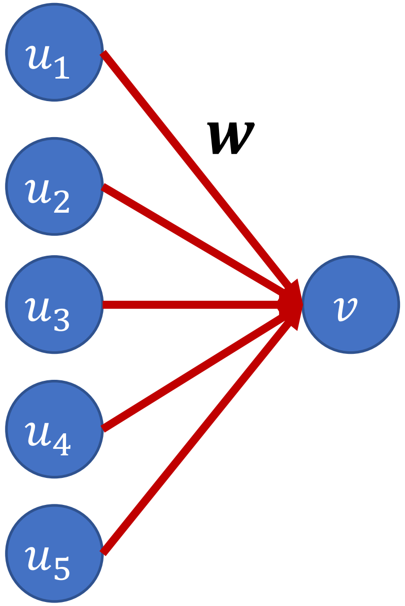
What if $w^T u = 0$ at first?

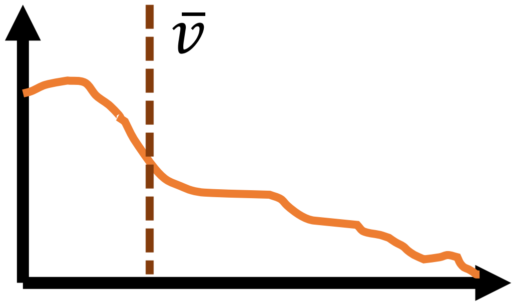
- What will happen to the weight w?
  - Nothing change.
  - Hebbian learning needs post-synaptic activity.

Extended Hebb's Rule / Covariance Rule

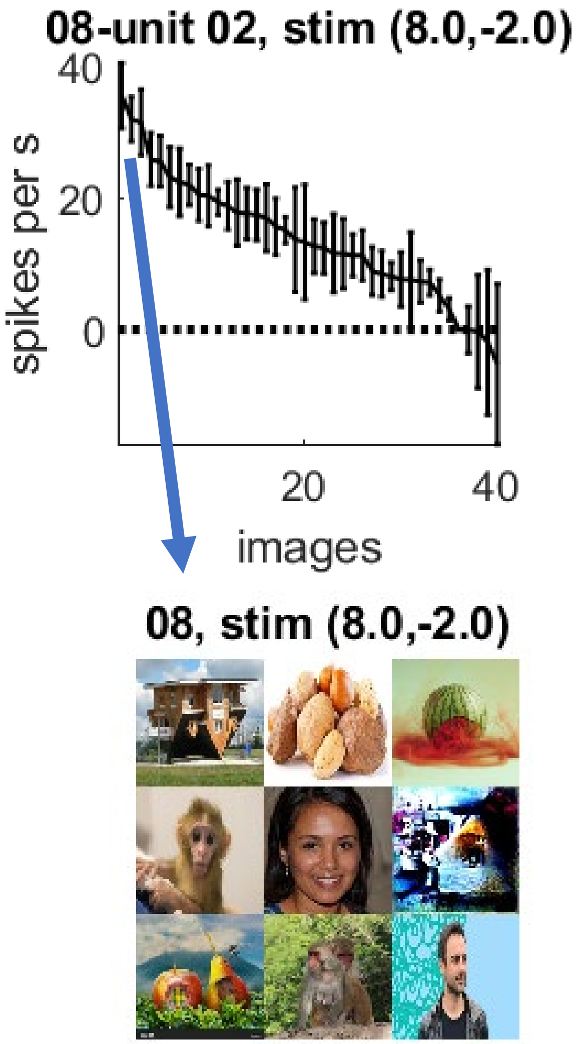
$$\tau \frac{dw}{dt} = u(v - \bar{v})$$

- Intuition
  - The u patterns evoking higher than average activity are strengthen. (preferred u patterns)
  - Other u patterns are suppressed.





v activity distrib.



Top preferred stimuli

Deriving the Covariance Rule

$$\tau \frac{dw}{dt} = u(v - \bar{v})$$

- Derive the covariance, $v = u^T w$

$$\tau \frac{dw}{dt} = u(u - \bar{u})^T w$$

$$\mathbb{E}[u(u - \bar{u})^T] = \mathbb{E}[(u - \bar{u})(u - \bar{u})^T] + \mathbb{E}[\bar{u}(u - \bar{u})]$$

$$= \mathbb{E}[(u - \bar{u})(u - \bar{u})^T] = \mathrm{cov}[u]$$

- Covariance based learning

$$\tau \frac{dw}{dt} = \mathrm{cov}[u] w$$

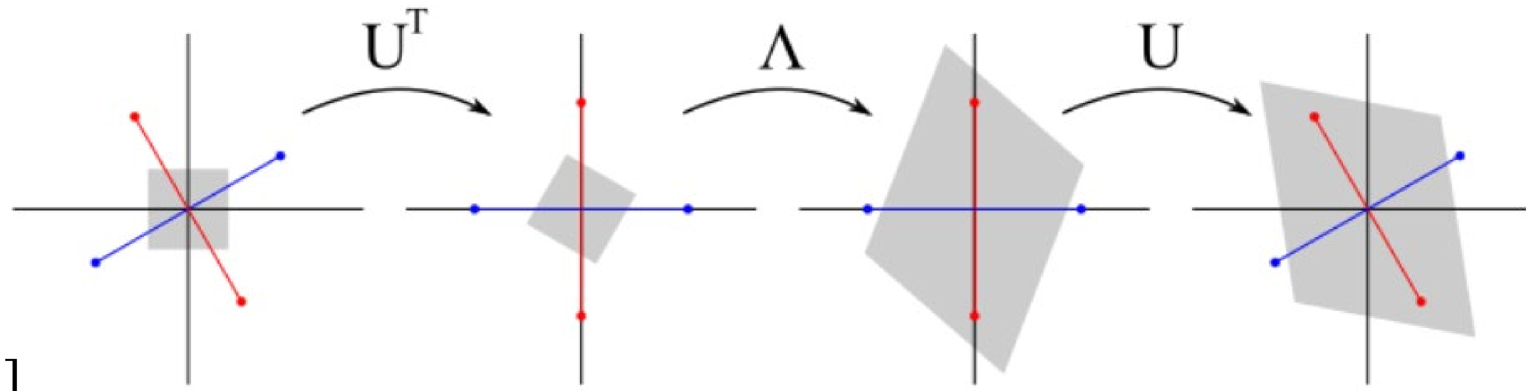
This derivation is approximated, u change faster than w, so we assume w stays the same in the averaging period.

$$\bar{v} = \frac{1}{T} \int_{t-T}^{t} v(t') dt'$$

$$= \frac{1}{T} \int_{t-T}^{t} w(t')^T u(t') dt'$$

$$\approx w(t)^T \frac{1}{T} \int_{t-T}^{t} u(t') dt'$$

Covariance Matrix



- For a data distribution $\{x_i\}$

$$\mathrm{cov}[x] = \mathbb{E}[(x-\bar{x})(x-\bar{x})^T]$$

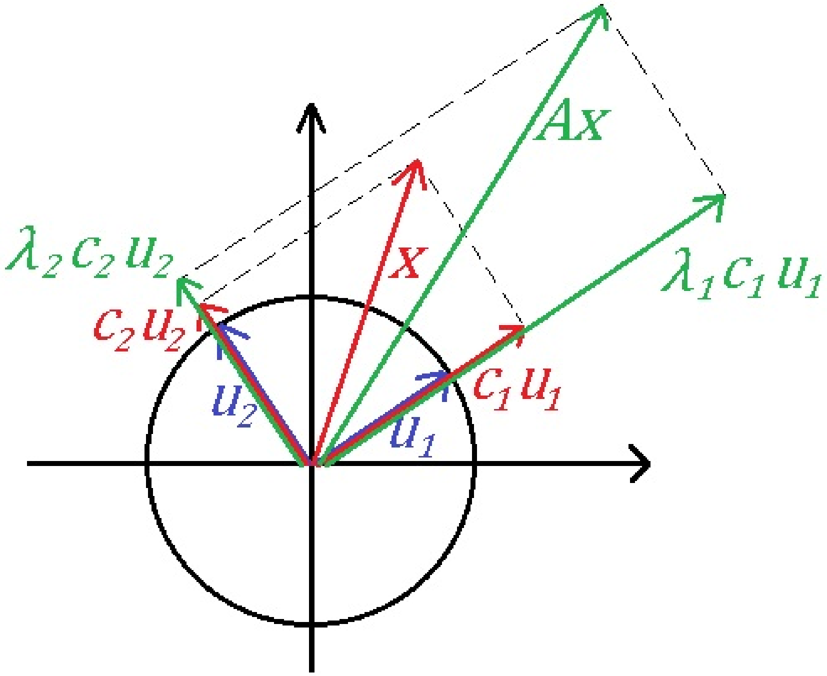
  - $\mathrm{cov}[x]_{ij} = \mathbb{E}[(x_i-\bar{x}_i)\left(x_j-\bar{x}_j\right)^T]$
  - $\mathrm{cov}[x] = (X-\bar{x})(X-\bar{x})^T$

  - What do we know about its eigenvalue and eigenvectors?
    - Real Symmetric matrix
    - Real eigenvalues.
    - Orthogonal eigenvectors
    - Eigenvalues are non-negative

  - What transform does this matrix represent ?

$$A = U\Lambda U^T$$

    - Scaling the n eigen dimension based on the n eigenvalues



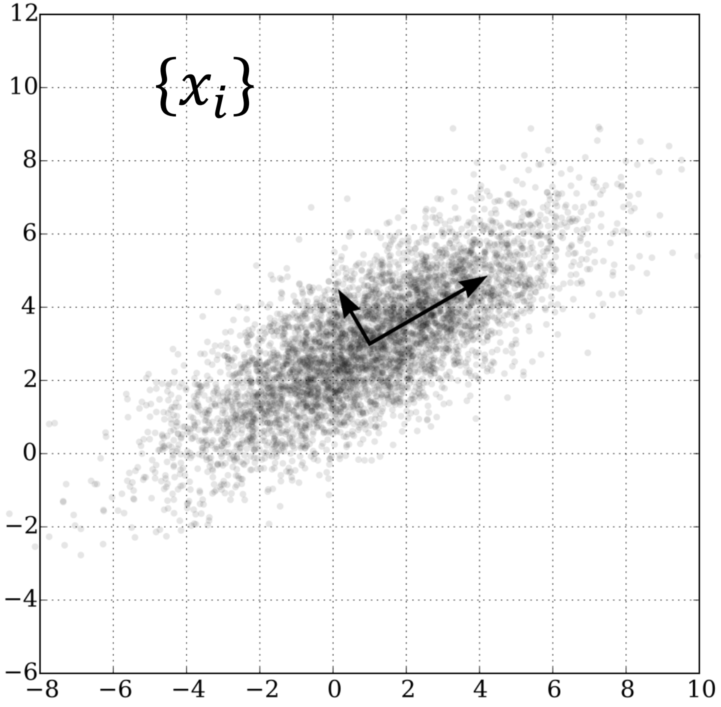
Covariance Matrix and PCA

- What's the connection between PCA and covariance matrix of data?

  - PCA correspond to eigenvectors of covariance matrix.

$$C = U\Lambda U^T$$

  - P orthogonal, Λ diagonal.
  - Columns of U are PC vectors, Λ_i is roughly explained variance



Eigenvalue and Power Iteration

- Dynamics of the weight vector

$$\tau \frac{dw}{dt} = \mathrm{cov}[u]w$$

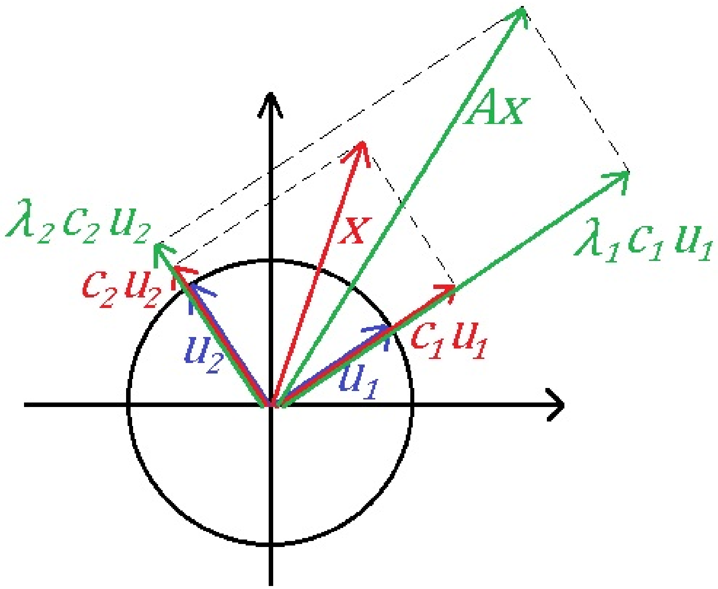
$$\tau \frac{dw}{dt} = U\Lambda U^T w$$

- Intuition
  - $\mathrm{cov}[u]$ amplify different eigen patterns in w based on their eigenvalues.
  - Top eigen vector $\boldsymbol{e_1}$ is amplified the most!

$$w(t) \propto \boldsymbol{e_1}, t \to \infty$$

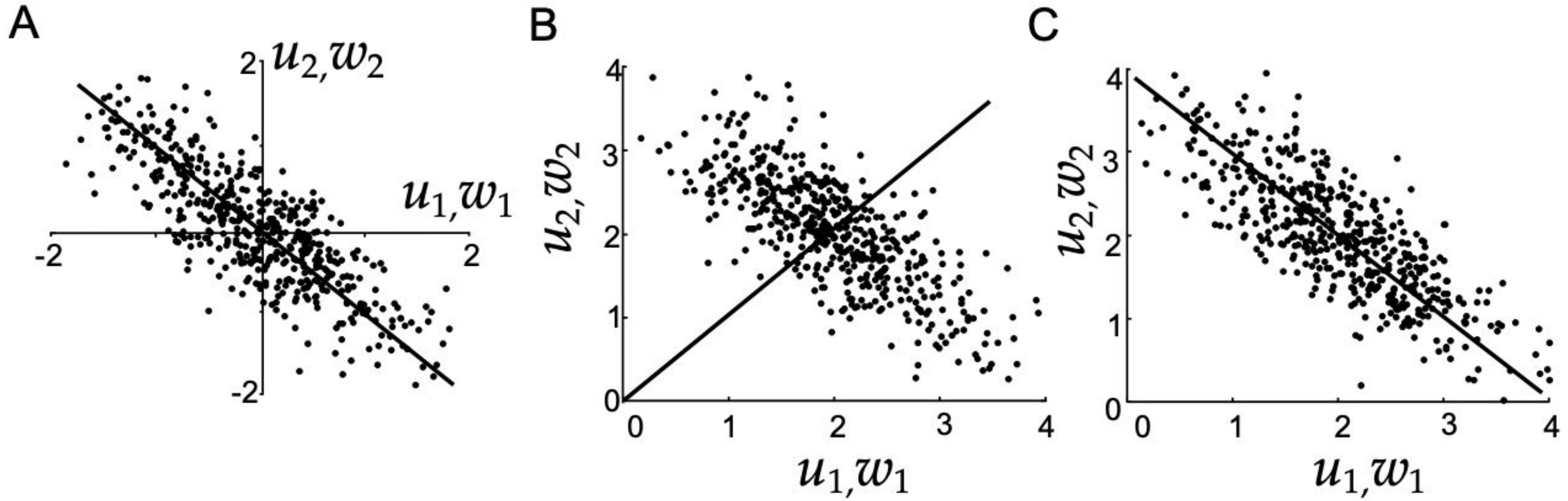
$$w(t) = \sum_i c_{0,i} e^{\lambda_i t} \boldsymbol{e_i}$$

  - Explode but in a specific direction!



Hebbian Learning Extracts PC1

- Single neuronal output represent the best 1d representation of its input!
  - $w \propto e_1, v = w^T u \propto e_1^T u$
  - Projection losing as little information as possible.



Oja's Rule

$$\tau \frac{dw}{dt} = vu - \alpha v^2 w$$

- Interpretation
  - vu is the Hebbian term.
  - $-\alpha v^2 w$ term decay / scale down each synapse by the same ratio.
    - Vector direction of w stay the same.

Weight Norm Analysis for Oja's Rule

- Weight norm analysis

$$\tau w^T \frac{dw}{dt} = w^{\mathrm{T}}(vu - \alpha v^2 w)$$

$$= v(w^T u) - \alpha v^2 w^T w$$
$$= v^2 - \alpha v^2 \|w\|^2$$
$$= v^2 (1 - \alpha \|w\|^2)$$

- Weight norm equation

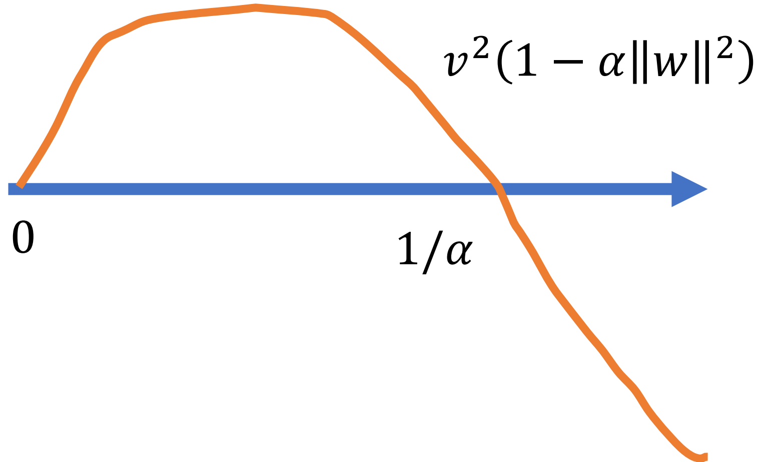
$$\frac{\tau}{2} \frac{d\|w\|^2}{dt} = v^2 (1 - \alpha \|w\|^2)$$

Oja's Rule and Synaptic Competition.

- What's the dynamic of $||w||$

$$\frac{\tau}{2}\frac{d||w||^2}{dt} = v^2(1 - \alpha||w||^2)$$
$$v = w^T u$$

  - Is it a 1d dynamic system?
  - Is it a linear system?
  - What's the flow on the line?
  - What's the stability of the fixed point?



Oja’s Rule and Synaptic Competition.

- Weight norm stays around the stable attractor :

$$\|w\| \approx \frac{1}{\sqrt{\alpha}}$$

- Interpretation
  - If some weights strengthen, others weaken.
  - Known as synaptic competition.

Summary of learning rules

Hebb's rule	$\tau_w \frac{d\bar{w}}{dt} = v\bar{u}$	Captures LTP	Weights explode
Covariance rule	$\tau_w \frac{d\bar{w}}{dt} = \bar{u}(v - < v >)$	Captures LTP & LTD	Weights explode
Oja's rule	$\tau_w \frac{d\bar{w}}{dt} = v\bar{u} - \alpha v^2 \bar{w}$	Captures LTP & LTD	Weights stable

General Takeaway

- By learning, information about input distribution $\{u_i\}$ are encoded in their weights w
  - PCA is one example.
  - Some suggests deep learning is doing the same thing.
- Original Hebbian learning is unstable, to make it homeostatic, synaptic competition is necessary.