Last time: Chaotic RNNs
This time: Low-rank RNNs (Mastrogiuseppe & Ostojic, 2018)

Observations:

1) Low-d dynamics in motor cortex during tasks:

2) Trained RNNs learn (approx.) low-rank updates to initial random connectivity (Schuessler et al., 2020)

\[ J = J_0 + P \]
Rank of a matrix:

\[ P_{ij} = \sum_{r=1}^{R} m_{r,i} n_{r,j} \]  \hspace{1cm} (2 vectors per rank)

\[ P = \sum_{r=1}^{R} m_{r} n_{r}^T \]

\[ P = M N^T \]  \hspace{1cm} \text{See SVD.}

RNN with low-rank component:

\[ \frac{dx_i}{dt} = -x_i + \sum_{j=1}^{N} J_{ij} \phi_j \]

\[ J = J_0 + P, \quad \text{is rank-1} \]

\[ a = \sum_{j=1}^{N} n_{ij} \phi_j \]

\[ \frac{dx_i}{dt} = -x_i + \sum_{j=1}^{N} J_{ij} \phi_j + M_i a \]
See e.g. Sussillo & Abbott 2009 (FORCE learning)

a "listens to" n direction and "talks to" m direction.

\( \mathbf{m}, \eta_r, \mathbf{I} \) are "special" directions in activity space.

Example: \( I = 0, R = 1 \).

\[ J_{ij} \sim N(0, g^2/N) \quad O(\sqrt{N}) \]

\[ P_{ij} = \frac{1}{N} m_i \eta_j \quad O(\sqrt{N}) \]

If \( \phi \) is independent (\& independent of \( J \))

\[ \text{Var} \left( \sum_{j=1}^{N} J_{ij} \phi_j \right) = N \text{Var} (J_{ij}) \text{Var} (\phi_j) \]

\[ \frac{g^2}{J_{ij}} \cdot N \cdot \text{Var} (\phi_j) \]

\[ \sum_{j=1}^{N} P_{ij} \phi_j = \frac{1}{N} m_i \sum_{j} \eta_j \phi_j \]
If $n$ and $\phi$ are independent, small.

If $\frac{1}{N}(n \cdot \phi) > 0$, $O(1)$.

How to make $\frac{1}{N}(n \cdot \phi) > 0$?

$\frac{1}{N}(m \cdot n) > 0$.

---

Diagram with axes $\frac{M \cdot n}{N}$ vs. $1$ showing regions labeled 1 to 4 with indications of fixed points (FP) and chaos transitions.
DMFT: \[ X = \frac{1}{N} \left( n - \phi \right) \]

Last class: \( X = 0 \).

Self-consistent equations for \( X, C(\tau) \):

\[ \frac{d x_i}{dt} = -x_i + \eta_i(t) \]

At steady state:

\[ \bar{x}_i = \langle \eta_i \rangle_t = m_i \cdot X \]

Let \( \Delta_x(\tau) = \langle (x_i(t) - \bar{x}_i)(x_i(t+\tau) - \bar{x}_i) \rangle_i \)

\[ m = \langle \eta \rangle_{t,i} = \langle m_i \rangle_X \]

\[ \Delta_x(\tau) = g^2 \Delta \phi(\tau) + \left( \langle m_i^2 \rangle - \langle m_i \rangle^2 \right) k^2 \]

\( C(\tau) \) from last class, extra variance from structure.

If statistics of \( m, n \) specified, can calculate \( m, \Delta_x(\tau) \) self-consistently.
For higher-rank networks, interactions described by

\[ x_1 \ldots x_r = \frac{1}{N}(n_1 - \phi) \]

Activity lies in subspace of \( m \ldots m_r \) and \( I \).

Example: \( m, n \) orthogonal but \( \frac{1}{N}(n \cdot I) > 0 \).

If \( \frac{1}{N}(m \cdot w) > 0 \),

can selectively respond to \( I \) & not other inputs.
Rank 2+: 1) Oscillations
2) Context-dependent responses