1 Convergence Properties of Oja's Rule

① small learning steps $\rightsquigarrow$ average over all patterns

$$\begin{aligned}
\Delta \mathrm{w}_j &= \frac{\epsilon}{p} \sum_{\alpha=1}^{p} \left\{ \sum_{k=1}^{N} \mathrm{w}_k \mathrm{x}_k^{(\alpha)} \mathrm{x}_j^{(\alpha)} - \mathrm{w}_j \sum_{k,l=1}^{N} \mathrm{w}_k \mathrm{w}_l \mathrm{x}_k^{(\alpha)} \mathrm{x}_l^{(\alpha)} \right\} \\
&= \epsilon \left\{ \sum_{k=1}^{N} \mathrm{w}_k C_{kj} - \mathrm{w}_j \sum_{k,l=1}^{N} \mathrm{w}_k \mathrm{w}_l C_{kl} \right\} \\
\Delta \underline{\mathbf{w}} &= \epsilon \Big\{ \underbrace{\underline{\mathbf{C}}\underline{\mathbf{w}}}_{\text{Hebbian rule}} - \underbrace{\underbrace{(\underline{\mathbf{w}}^T \underline{\mathbf{C}}\underline{\mathbf{w}})}_{\text{always positive}} \underline{\mathbf{w}}}_{\text{decay term}} \Big\}
\end{aligned} \tag{1}$$

② stationary states $\underline{\mathbf{w}}^*$ of Oja's rule
$\widehat{=}$ normalized eigenvectors $\underline{\mathbf{e}}_j$ of the correlation matrix $\underline{\mathbf{C}}$

Proof:

stationary state: $\Delta \underline{\mathbf{w}} \overset{!}{=} \underline{\mathbf{0}}$

ansatz: $\underline{\mathbf{w}}^* = C \underline{\mathbf{e}}_j \sim \underline{\mathbf{e}}_j$

insertion into Oja's rule:

$$\begin{aligned}
&\Delta \underline{\mathbf{w}} = \epsilon \Big\{ C \lambda_j \underline{\mathbf{e}}_j - C^3 \lambda_j \underline{\mathbf{e}}_j \Big\} \overset{!}{=} O \\
&\Rightarrow C^2 = 1 \\
&\Rightarrow \underline{\mathbf{w}}^* = \pm \underline{\mathbf{e}}_j
\end{aligned} \tag{2}$$

③ The stationary state $\underline{\mathbf{w}}^* = \underline{\mathbf{e}}_j$ is stable if and only if $\underline{\mathbf{e}}_j = \underline{\mathbf{e}}_1$, i.e. if $\underline{\mathbf{e}}_j$ is the eigenvactor with the largest eigenvalue.

Proof: linear stability analysis

$\lambda_1 > \lambda_2 > \ldots > \lambda_N$ eigenvalues of $\underline{\mathbf{C}}$

$\underline{\mathbf{w}} = \underline{\mathbf{e}}_j + \underline{\eta} \leftarrow$ small deviation from the stationary state

$$\begin{aligned}
\Delta \underline{\mathbf{w}} = \quad \Delta \underline{\eta} &= \epsilon \underline{\mathbf{C}} (\underline{\mathbf{e}}_j + \underline{\eta}) - \epsilon \Big\{ (\underline{\mathbf{e}}_j + \underline{\eta})^T \underline{\mathbf{C}} (\underline{\mathbf{e}}_j + \underline{\eta}) \Big\} (\underline{\mathbf{e}}_j + \underline{\eta}) \\
&= \epsilon \Big\{ \lambda_j \underline{\mathbf{e}}_j + \underline{\mathbf{C}} \underline{\eta} - \lambda_j \underline{\mathbf{e}}_j - \lambda_j \underline{\eta} - 2\lambda_j (\underline{\mathbf{e}}_j^T \underline{\eta})(\underline{\mathbf{e}}_j + \underline{\eta}) - (\underline{\eta}^T \underline{\mathbf{C}} \underline{\eta})(\underline{\mathbf{e}}_j + \underline{\eta}) \Big\} \\
\Delta \eta &= \epsilon \Big\{ - 2\lambda_j (\underline{\mathbf{e}}_j^T \underline{\eta}) \underline{\mathbf{e}}_j + \underline{\mathbf{C}} \underline{\eta} - \lambda_j \underline{\eta} \Big\} + \underbrace{O_{(\underline{\eta}^2)}}_{\text{discarded for } |\underline{\eta}| \to 0}
\end{aligned} \tag{3}$$

Projection onto the eigenvectors $\underline{\mathbf{e}}_k$:

$$\underline{\mathbf{e}}_k^T \Delta \underline{\eta} = \epsilon \Big\{ (\lambda_k - \lambda_j) \underline{\mathbf{e}}_k^T \underline{\eta} - 2\lambda_k (\underline{\mathbf{e}}_k^T \underline{\eta}) \underbrace{\delta_{kj}}_{\substack{\text{Kronecker-}\\ \text{Delta}}} \Big\} \tag{4}$$

<u>case I:</u> $k = j$

$$\begin{aligned} \underline{\mathbf{e}}_k^T \Delta \underline{\eta} &= \underbrace{-2\epsilon\lambda_k}_{\substack{\text{always}\\ \text{negative}}} \underline{\mathbf{e}}_k^T \underline{\eta} \\ &\Rightarrow \underline{\mathbf{e}}_k^T \underline{\eta} \to 0 \end{aligned} \tag{5}$$

<u>case II:</u> $k \neq j$

$$\begin{aligned} \underline{\mathbf{e}}_k^T \Delta \underline{\eta} = \quad & \underbrace{\epsilon(\lambda_k - \lambda_j)}_{\substack{\text{factor is negative}\\ \text{only if } \lambda_j \text{ is}\\ \text{the largest eigenvalue}}} \quad \underline{\mathbf{e}}_k^T \underline{\eta} \\ \Rightarrow \ \text{if } \lambda_j = \lambda_1, \ \text{then } & \underline{\mathbf{e}}_k^T \underline{\eta} \to 0 \end{aligned} \tag{6}$$