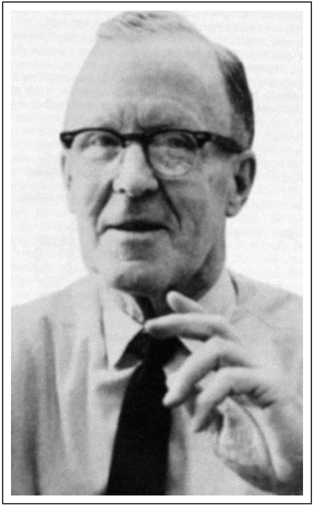
Hebbian Learning



Donald Hebb (1904–1985)

Hebbian Learning

"When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased." *(Donald Hebb, 1949)*

Hebbian Learning

"When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased." *(Donald Hebb, 1949)*

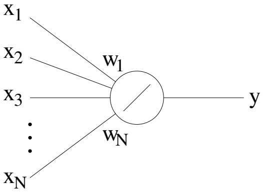
"Fire together, wire together"

Hebbian Learning

"When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased." *(Donald Hebb, 1949)*

"Fire together, wire together"

Simple model of Hebbian learning:



$$\Delta w_i = \eta \frac{1}{M} \sum_\mu y^\mu x_i^\mu \quad (1)$$

$$= \eta \langle y^\mu x_i^\mu \rangle_\mu \quad (2)$$

$$\iff \Delta \mathbf{w} = \eta \langle y^\mu \mathbf{x}^\mu \rangle_\mu \,. \quad (3)$$

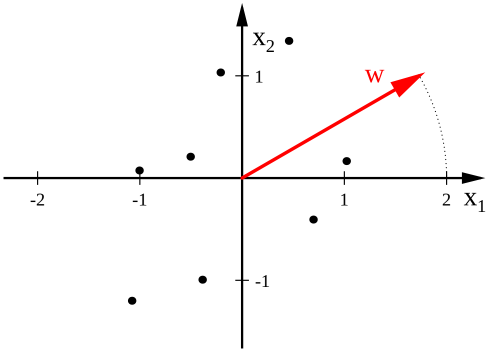
In the following we will drop μ for convenience.

Linear Model Neuron

A linear model neuron is described by:

$$y = \sum_{i=1}^{N} w_i x_i = \mathbf{w}^T \mathbf{x}\,. \qquad (4)$$

This corresponds to a projection of the data onto the axis given by **w** scaled with $\| \mathbf{w} \|$.

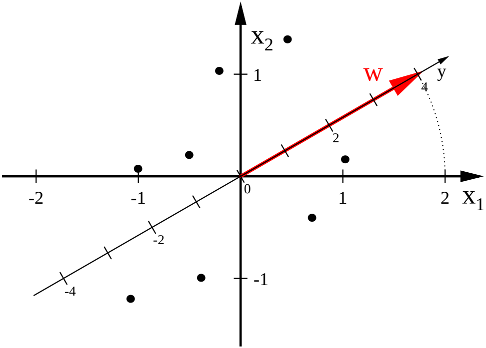


Linear Model Neuron

A linear model neuron is described by:

$$y = \sum_{i=1}^{N} w_i x_i = \mathbf{w}^T \mathbf{x} \,. \tag{4}$$

This corresponds to a projection of the data onto the axis given by **w** scaled with $\| \mathbf{w} \|$.

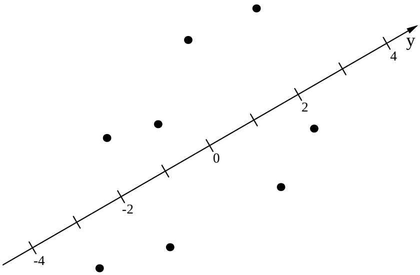


Linear Model Neuron

A linear model neuron is described by:

$$y = \sum_{i=1}^{N} w_i x_i = \mathbf{w}^T \mathbf{x}\,. \qquad (4)$$

This corresponds to a projection of the data onto the axis given by **w** scaled with $\| \mathbf{w} \|$.

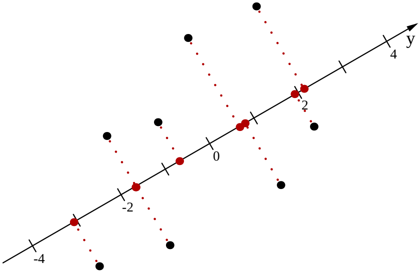


Linear Model Neuron

A linear model neuron is described by:

$$y = \sum_{i=1}^{N} w_i x_i = \mathbf{w}^T \mathbf{x}\,. \tag{4}$$

This corresponds to a projection of the data onto the axis given by $\mathbf{w}$ scaled with $\| \mathbf{w} \|$.

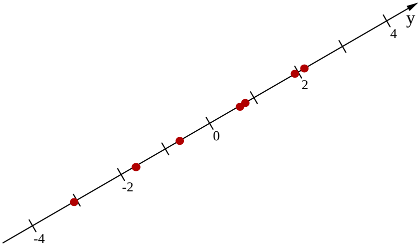


Linear Model Neuron

A linear model neuron is described by:

$$y = \sum_{i=1}^{N} w_i x_i = \mathbf{w}^T \mathbf{x}\,. \tag{4}$$

This corresponds to a projection of the data onto the axis given by **w** scaled with $\| \mathbf{w} \|$.



Hebbian Learning with a Linear Model Neuron

$$\Delta \mathbf{w} \stackrel{(3)}{=} \eta \langle y \mathbf{x} \rangle \quad \text{(Hebbian learning)}$$

$$y \stackrel{(4)}{=} \mathbf{w}^T \mathbf{x}\,, \quad \text{(linear neuron)}$$

Hebbian Learning with a Linear Model Neuron

$$\Delta \mathbf{w} \stackrel{(3)}{=} \eta \langle y \mathbf{x} \rangle \quad \text{(Hebbian learning)}$$

$$y \stackrel{(4)}{=} \mathbf{w}^T \mathbf{x}\,, \quad \text{(linear neuron)}$$

$$\begin{aligned}
\Delta \mathbf{w} &\stackrel{(3)}{=} \eta \langle y \mathbf{x} \rangle && (5)\\
&\stackrel{(4)}{=} \eta \left\langle (\mathbf{w}^T \mathbf{x}) \mathbf{x} \right\rangle && (6)\\
&= \eta \left\langle \mathbf{x} (\mathbf{x}^T \mathbf{w}) \right\rangle && (7)\\
&= \eta \left\langle \mathbf{x} \mathbf{x}^T \right\rangle \mathbf{w} && (8)
\end{aligned}$$

Hebbian Learning with a Linear Model Neuron

$$\begin{aligned} \Delta\mathbf{w} &\overset{(3)}{=} \eta\langle y\mathbf{x}\rangle \quad \text{(Hebbian learning)} \\ y &\overset{(4)}{=} \mathbf{w}^T\mathbf{x}\,, \quad \text{(linear neuron)} \end{aligned}$$

$$\Delta\mathbf{w} \overset{(3)}{=} \eta\langle y\mathbf{x}\rangle \tag{5}$$

$$\overset{(4)}{=} \eta\left\langle(\mathbf{w}^T\mathbf{x})\mathbf{x}\right\rangle \tag{6}$$

$$= \eta\left\langle\mathbf{x}(\mathbf{x}^T\mathbf{w})\right\rangle \tag{7}$$

$$= \eta\left\langle\mathbf{x}\mathbf{x}^T\right\rangle\mathbf{w} \tag{8}$$

$$\Longrightarrow \quad \mathbf{w}^{t+1} = \mathbf{w}^t + \eta\left\langle\mathbf{x}\mathbf{x}^T\right\rangle\mathbf{w}^t \tag{9}$$

$$= \left(\mathbf{I} + \eta\left\langle\mathbf{x}\mathbf{x}^T\right\rangle\right)\mathbf{w}^t \tag{10}$$

Hebbian Learning with a Linear Model Neuron

$$\begin{aligned} \Delta \mathbf{w} &\overset{(3)}{=} \eta \langle y \mathbf{x} \rangle \quad \text{(Hebbian learning)} \\ y &\overset{(4)}{=} \mathbf{w}^T \mathbf{x}\,, \quad \text{(linear neuron)} \end{aligned}$$

$$\begin{aligned} \Delta \mathbf{w} &\overset{(3)}{=} \eta \langle y \mathbf{x} \rangle && (5) \\ &\overset{(4)}{=} \eta \left\langle (\mathbf{w}^T \mathbf{x}) \mathbf{x} \right\rangle && (6) \\ &= \eta \left\langle \mathbf{x} (\mathbf{x}^T \mathbf{w}) \right\rangle && (7) \\ &= \eta \left\langle \mathbf{x}\mathbf{x}^T \right\rangle \mathbf{w} && (8) \\ \Longrightarrow \quad \mathbf{w}^{t+1} &= \mathbf{w}^t + \eta \left\langle \mathbf{x}\mathbf{x}^T \right\rangle \mathbf{w}^t && (9) \\ &= \left(\mathbf{I} + \eta \left\langle \mathbf{x}\mathbf{x}^T \right\rangle \right) \mathbf{w}^t && (10) \end{aligned}$$

Hebbian learning with a linear model neuron can be interpreted as an iterated multiplication of the weight vector $\mathbf{w}$ with matrix $(\mathbf{I} + \eta \langle \mathbf{x}\mathbf{x}^T \rangle)$.

Hebbian Learning with a Linear Model Neuron

$$\begin{aligned} \Delta \mathbf{w} &\overset{(3)}{=} \eta \langle y \mathbf{x} \rangle \quad \text{(Hebbian learning)} \\ y &\overset{(4)}{=} \mathbf{w}^T \mathbf{x}\,, \quad \text{(linear neuron)} \end{aligned}$$

$$\begin{aligned} \Delta \mathbf{w} &\overset{(3)}{=} \eta \langle y \mathbf{x} \rangle && (5) \\ &\overset{(4)}{=} \eta \left\langle (\mathbf{w}^T \mathbf{x}) \mathbf{x} \right\rangle && (6) \\ &= \eta \left\langle \mathbf{x} (\mathbf{x}^T \mathbf{w}) \right\rangle && (7) \\ &= \eta \left\langle \mathbf{x} \mathbf{x}^T \right\rangle \mathbf{w} && (8) \\ \Longrightarrow \quad \mathbf{w}^{t+1} &= \mathbf{w}^t + \eta \left\langle \mathbf{x} \mathbf{x}^T \right\rangle \mathbf{w}^t && (9) \\ &= \left(\mathbf{I} + \eta \left\langle \mathbf{x} \mathbf{x}^T \right\rangle \right) \mathbf{w}^t && (10) \end{aligned}$$

Hebbian learning with a linear model neuron can be interpreted as an iterated multiplication of the weight vector $\mathbf{w}$ with matrix $(\mathbf{I} + \eta \langle \mathbf{x}\mathbf{x}^T \rangle)$.

The 'sign' of the input vectors $\mathbf{x}$ is irrelevant for the learning.

Illustration of Hebbian Learning

$$\Delta \mathbf{w} \overset{(6)}{=} \eta \left\langle (\mathbf{w}^T \mathbf{x})\mathbf{x} \right\rangle \quad \text{(Hebbian learning for a linear unit)}$$

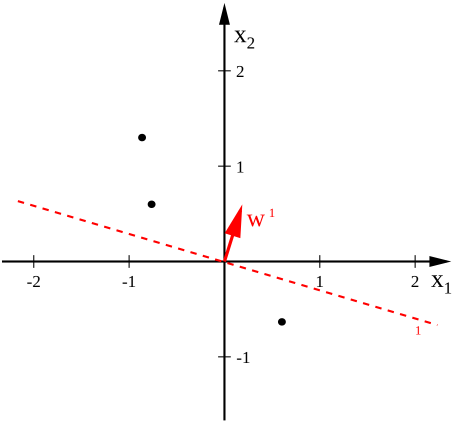


Illustration of Hebbian Learning

$$\Delta \mathbf{w} \stackrel{(6)}{=} \eta \left\langle (\mathbf{w}^T \mathbf{x})\mathbf{x} \right\rangle \quad \text{(Hebbian learning for a linear unit)}$$

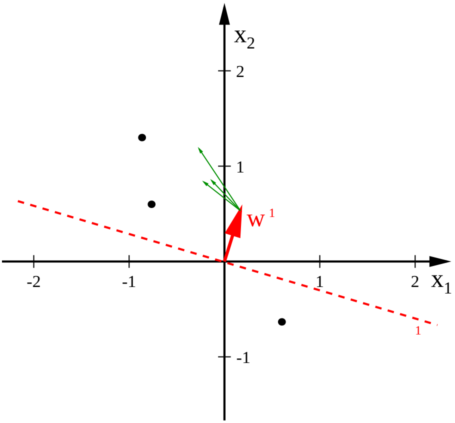


Illustration of Hebbian Learning

$$\Delta \mathbf{w} \stackrel{(6)}{=} \eta \left\langle (\mathbf{w}^T \mathbf{x})\mathbf{x} \right\rangle \quad \text{(Hebbian learning for a linear unit)}$$

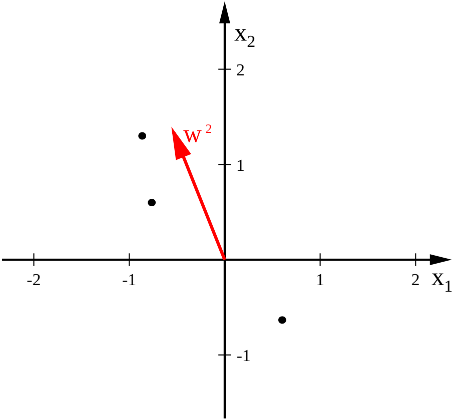


Illustration of Hebbian Learning

$$\Delta \mathbf{w} \overset{(6)}{=} \eta \left\langle (\mathbf{w}^T \mathbf{x})\mathbf{x} \right\rangle \quad \text{(Hebbian learning for a linear unit)}$$

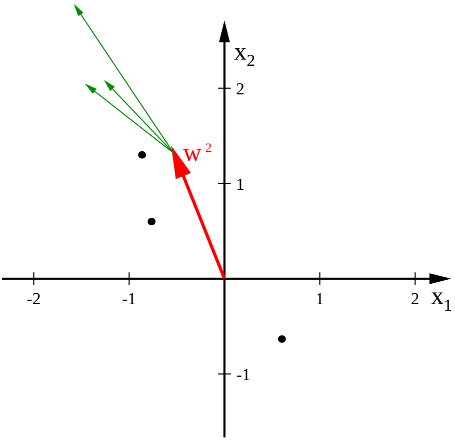


Illustration of Hebbian Learning

$$\Delta\mathbf{w} \overset{(6)}{=} \eta\left\langle(\mathbf{w}^T\mathbf{x})\mathbf{x}\right\rangle \quad \text{(Hebbian learning for a linear unit)}$$

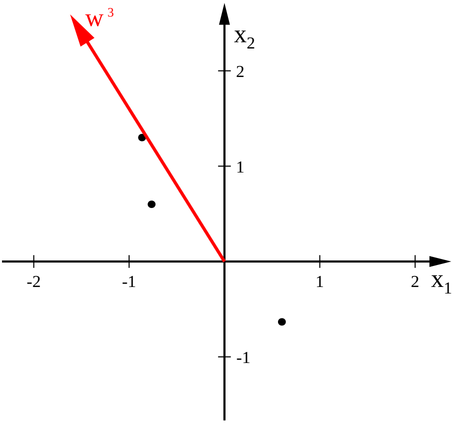


Illustration of Hebbian Learning

$$\Delta \mathbf{w} \overset{(6)}{=} \eta \left\langle (\mathbf{w}^T \mathbf{x}) \mathbf{x} \right\rangle \quad \text{(Hebbian learning for a linear unit)}$$

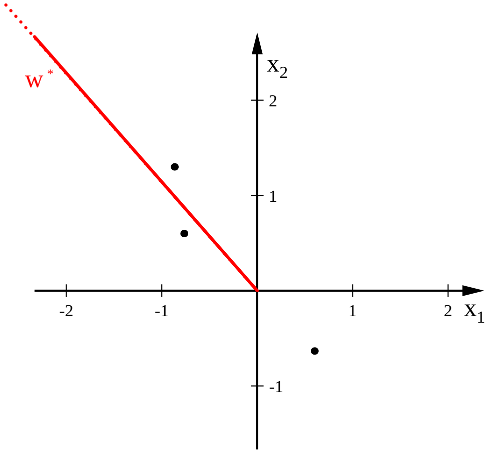
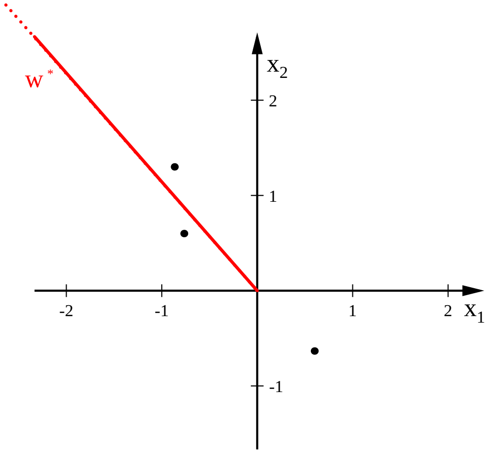


Illustration of Hebbian Learning

$$\Delta \mathbf{w} \overset{(6)}{=} \eta \left\langle (\mathbf{w}^T \mathbf{x}) \mathbf{x} \right\rangle \quad \text{(Hebbian learning for a linear unit)}$$



Problem: The weights grow unlimited.

Explicit Normalization of the Weight Vectors

First apply the learning rule,

$$\tilde{w}_i^{t+1} = w_i^t + \eta f_i \quad \text{with, e.g.,} f_i = y x_i \, . \tag{11}$$

Then normalize the weight vector to length one,

$$w_i^{t+1} = \frac{\tilde{w}_i^{t+1}}{\sqrt{\sum_j \left(\tilde{w}_j^{t+1}\right)^2}} \, . \tag{12}$$

We can verify that

$$\sqrt{\sum_i \left(w_i^{t+1}\right)^2} \overset{(12)}{=} \sqrt{\sum_i \left(\frac{\tilde{w}_i^{t+1}}{\sqrt{\sum_j \left(\tilde{w}_j^{t+1}\right)^2}} \right)^2} = \sqrt{\frac{\sum_i \left(\tilde{w}_i^{t+1}\right)^2}{\sum_j \left(\tilde{w}_j^{t+1}\right)^2}} = 1 \, . \tag{13}$$

Problem: Explicit normalization is expensive and biologically implausible.

Implicit Normalization of the Weight Vector

Explicit normalization yielded

$$w_i^{t+1}(\eta) \overset{(12)}{=} \frac{w_i^t + \eta f_i}{\sqrt{\sum_j \left(w_j^t + \eta f_j\right)^2}} \,.$$

Implicit Normalization of the Weight Vector

Explicit normalization yielded

$$w_i^{t+1}(\eta) \quad \overset{(12)}{=} \quad \frac{w_i^t + \eta f_i}{\sqrt{\sum_j \left(w_j^t + \eta f_j\right)^2}} \, .$$

To optain a simpler/more plausible normalization rule, we compute the Taylor-expansion of w_i^{t+1} in η at $\eta' = 0$ under the assumption that $\mathbf{w}^t$ is normalized, i.e. $\sum_j (w_j^t)^2 = 1$,

$$w_i^{t+1}(\eta) \quad \approx \quad w_i^{t+1}(\eta' = 0) + \eta \left. \frac{\partial w_i^{t+1}(\eta')}{\partial \eta'} \right|_{\eta'=0} \tag{14}$$

Implicit Normalization of the Weight Vector

Explicit normalization yielded

$$w_i^{t+1}(\eta) \quad \overset{(12)}{=} \quad \frac{w_i^t + \eta f_i}{\sqrt{\sum_j \left(w_j^t + \eta f_j \right)^2}} .$$

To optain a simpler/more plausible normalization rule, we compute the Taylor-expansion of w_i^{t+1} in η at $\eta' = 0$ under the assumption that $\mathbf{w}^t$ is normalized, i.e. $\sum_j (w_j^t)^2 = 1$,

$$w_i^{t+1}(\eta) \quad \approx \quad w_i^{t+1}(\eta' = 0) + \eta \left. \frac{\partial w_i^{t+1}(\eta')}{\partial \eta'} \right|_{\eta'=0} \tag{14}$$

$$= \quad w_i^{t+1}(\eta' = 0) + \eta \left. \left(\frac{f_i \sqrt{\sum_j \left(w_j^t + \eta' f_j \right)^2} - (w_i^t + \eta' f_i) \frac{\sum_j f_j \left(w_j^t + \eta' f_j \right)}{\sqrt{\sum_j \left(w_j^t + \eta' f_j \right)^2}}}{\sum_j \left(w_j^t + \eta' f_j \right)^2} \right) \right|_{\eta'=0} \tag{15}$$

Implicit Normalization of the Weight Vector

Explicit normalization yielded

$$w_i^{t+1}(\eta) \stackrel{(12)}{=} \frac{w_i^t + \eta f_i}{\sqrt{\sum_j \left(w_j^t + \eta f_j\right)^2}} .$$

To optain a simpler/more plausible normalization rule, we compute the Taylor-expansion of w_i^{t+1} in η at $\eta' = 0$ under the assumption that $\mathbf{w}^t$ is normalized, i.e. $\sum_j (w_j^t)^2 = 1$,

$$w_i^{t+1}(\eta) \approx w_i^{t+1}(\eta' = 0) + \eta \left. \frac{\partial w_i^{t+1}(\eta')}{\partial \eta'} \right|_{\eta'=0} \tag{14}$$

$$= w_i^{t+1}(\eta' = 0) + \eta \left. \left(\frac{f_i \sqrt{\sum_j \left(w_j^t + \eta' f_j\right)^2} - (w_i^t + \eta' f_i) \frac{\sum_j f_j \left(w_j^t + \eta' f_j\right)}{\sqrt{\sum_j \left(w_j^t + \eta' f_j\right)^2}}}{\sum_j \left(w_j^t + \eta' f_j\right)^2} \right) \right|_{\eta'=0} \tag{15}$$

$$= w_i^t + \eta \left(f_i - w_i^t \sum_j f_j w_j^t \right) . \tag{16}$$

Implicit Normalization of the Weight Vector

Taylor-expansion of the explicit normalization rule resulted in

$$w_i^{t+1} \quad \overset{(16)}{\approx} \quad w_i^t + \eta \left(f_i - w_i^t \sum_j f_j w_j^t \right) \ .$$

Implicit Normalization of the Weight Vector

Taylor-expansion of the explicit normalization rule resulted in

$$w_i^{t+1} \quad \overset{(16)}{\approx} \quad w_i^t + \eta \left(f_i - w_i^t \sum_j f_j w_j^t \right) .$$

For Hebbian learning, i.e. with $f_i = yx_i$ and $y = \sum_j x_j w_j$, this becomes

$$\Delta w_i \quad \overset{(16)}{=} \quad \eta \left(f_i - w_i \sum_j f_j w_j \right) \tag{17}$$

Implicit Normalization of the Weight Vector

Taylor-expansion of the explicit normalization rule resulted in

$$w_i^{t+1} \stackrel{(16)}{\approx} w_i^t + \eta \left(f_i - w_i^t \sum_j f_j w_j^t \right) .$$

For Hebbian learning, i.e. with $f_i = yx_i$ and $y = \sum_j x_j w_j$, this becomes

$$\Delta w_i \stackrel{(16)}{=} \eta \left(f_i - w_i \sum_j f_j w_j \right) \tag{17}$$

$$= \eta \left(yx_i - w_i y \underbrace{\sum_j x_j w_j}_{=y} \right) \tag{18}$$

Implicit Normalization of the Weight Vector

Taylor-expansion of the explicit normalization rule resulted in

$$w_i^{t+1} \overset{(16)}{\approx} w_i^t + \eta \left(f_i - w_i^t \sum_j f_j w_j^t \right) .$$

For Hebbian learning, i.e. with $f_i = yx_i$ and $y = \sum_j x_j w_j$, this becomes

$$\Delta w_i \overset{(16)}{=} \eta \left(f_i - w_i \sum_j f_j w_j \right) \tag{17}$$

$$= \eta \left(yx_i - w_i y \underbrace{\sum_j x_j w_j}_{=y} \right) \tag{18}$$

$$= \eta \left(yx_i - y^2 w_i \right) \quad \forall i \tag{19}$$

Implicit Normalization of the Weight Vector

Taylor-expansion of the explicit normalization rule resulted in

$$w_i^{t+1} \quad \overset{(16)}{\approx} \quad w_i^t + \eta \left(f_i - w_i^t \sum_j f_j w_j^t \right) .$$

For Hebbian learning, i.e. with $f_i = yx_i$ and $y = \sum_j x_j w_j$, this becomes

$$\Delta w_i \quad \overset{(16)}{=} \quad \eta \left(f_i - w_i \sum_j f_j w_j \right) \tag{17}$$

$$= \quad \eta \left(yx_i - w_i y \underbrace{\sum_j x_j w_j}_{=y} \right) \tag{18}$$

$$= \quad \eta \left(yx_i - y^2 w_i \right) \quad \forall i \tag{19}$$

$$\Longleftrightarrow \quad \Delta \mathbf{w} \quad = \quad \eta \left(y\mathbf{x} - y^2 \mathbf{w} \right) . \quad \text{(Oja's rule)} \tag{20}$$

Implicit Normalization of the Weight Vector

Taylor-expansion of the explicit normalization rule resulted in

$$w_i^{t+1} \overset{(16)}{\approx} w_i^t + \eta \left(f_i - w_i^t \sum_j f_j w_j^t \right) .$$

For Hebbian learning, i.e. with $f_i = y x_i$ and $y = \sum_j x_j w_j$, this becomes

$$\Delta w_i \overset{(16)}{=} \eta \left(f_i - w_i \sum_j f_j w_j \right) \quad (17)$$

$$= \eta \left(y x_i - w_i y \underbrace{\sum_j x_j w_j}_{=y} \right) \quad (18)$$

$$= \eta \left(y x_i - y^2 w_i \right) \quad \forall i \quad (19)$$

$$\Longleftrightarrow \quad \Delta \mathbf{w} = \eta \left(y \mathbf{x} - y^2 \mathbf{w} \right) . \quad \text{(Oja's rule)} \quad (20)$$

This rule is simpler, e.g. Δw_i does not depend on w_j anymore.

Oja's Rule

Implicit normalization yielded

$$\Delta \mathbf{w} \overset{(20)}{=} \eta \left(y\mathbf{x} - y^2 \mathbf{w} \right). \quad \text{(Oja's rule)}$$

Oja's Rule

Implicit normalization yielded

$$\Delta \mathbf{w} \overset{(20)}{=} \eta \left(y\mathbf{x} - y^2 \mathbf{w} \right). \quad \text{(Oja's rule)}$$

For small y we have

Oja's Rule

Implicit normalization yielded

$$\Delta \mathbf{w} \overset{(20)}{=} \eta \left(y\mathbf{x} - y^2 \mathbf{w} \right). \quad \text{(Oja's rule)}$$

For small y we have

$$\Delta \mathbf{w} \approx \eta y \mathbf{x}, \tag{21}$$

Oja's Rule

Implicit normalization yielded

$$\Delta \mathbf{w} \overset{(20)}{=} \eta \left(y\mathbf{x} - y^2 \mathbf{w} \right) . \quad \text{(Oja's rule)}$$

For small y we have

$$\Delta \mathbf{w} \approx \eta y \mathbf{x} \, , \tag{21}$$

which corresponds to Hebb's rule.

Oja's Rule

Implicit normalization yielded

$$\Delta \mathbf{w} \overset{(20)}{=} \eta \left(y\mathbf{x} - y^2 \mathbf{w} \right). \quad \text{(Oja's rule)}$$

For small y we have

$$\Delta \mathbf{w} \approx \eta y \mathbf{x}, \tag{21}$$

which corresponds to Hebb's rule.

For large y we have

Oja's Rule

Implicit normalization yielded

$$\Delta \mathbf{w} \quad \stackrel{(20)}{=} \quad \eta \left(y\mathbf{x} - y^2 \mathbf{w} \right). \quad \text{(Oja's rule)}$$

For small y we have

$$\Delta \mathbf{w} \approx \eta y \mathbf{x}\,, \tag{21}$$

which corresponds to Hebb's rule.

For large y we have

$$\Delta \mathbf{w} \approx -\eta y^2 \mathbf{w}\,, \tag{22}$$

Oja's Rule

Implicit normalization yielded

$$\Delta \mathbf{w} \overset{(20)}{=} \eta \left(y\mathbf{x} - y^2 \mathbf{w} \right). \quad \text{(Oja's rule)}$$

For small y we have

$$\Delta \mathbf{w} \approx \eta y \mathbf{x}\,, \tag{21}$$

which corresponds to Hebb's rule.

For large y we have

$$\Delta \mathbf{w} \approx -\eta y^2 \mathbf{w}\,, \tag{22}$$

which results in a shortening of the weight vector.

Oja's Rule

Implicit normalization yielded

$$\Delta \mathbf{w} \overset{(20)}{=} \eta \left(y\mathbf{x} - y^2 \mathbf{w} \right) . \quad \text{(Oja's rule)}$$

For small y we have

$$\Delta \mathbf{w} \approx \eta y \mathbf{x} , \tag{21}$$

which corresponds to Hebb's rule.

For large y we have

$$\Delta \mathbf{w} \approx -\eta y^2 \mathbf{w} , \tag{22}$$

which results in a shortening of the weight vector.

The second term only affects the length but not the direction of the weight vector. It therefore limits the length of the weight vector without affecting the qualitative behavior of the learning rule.

Oja's Rule

Implicit normalization yielded

$$\Delta\mathbf{w} \stackrel{(20)}{=} \eta\left(y\mathbf{x} - y^2\mathbf{w}\right). \quad \text{(Oja's rule)}$$

For small y we have

$$\Delta\mathbf{w} \approx \eta y\mathbf{x}\,, \tag{21}$$

which corresponds to Hebb's rule.

For large y we have

$$\Delta\mathbf{w} \approx -\eta y^2\mathbf{w}\,, \tag{22}$$

which results in a shortening of the weight vector.

The second term only affects the length but not the direction of the weight vector. It therefore limits the length of the weight vector without affecting the qualitative behavior of the learning rule.

Question: What will be the final length of the weight vector?

Oja's Rule

Implicit normalization yielded

$$\Delta \mathbf{w} \overset{(20)}{=} \eta \left(y\mathbf{x} - y^2 \mathbf{w} \right) . \quad \text{(Oja's rule)}$$

For small y we have

$$\Delta \mathbf{w} \approx \eta y \mathbf{x} \,, \tag{21}$$

which corresponds to Hebb's rule.

For large y we have

$$\Delta \mathbf{w} \approx -\eta y^2 \mathbf{w} \,, \tag{22}$$

which results in a shortening of the weight vector.

The second term only affects the length but not the direction of the weight vector. It therefore limits the length of the weight vector without affecting the qualitative behavior of the learning rule.

Question: What will be the final length of the weight vector?

Question: How will the weight vector develop under Oja's rule in general?

Linear Stability Analysis

$$y \stackrel{(4)}{=} \mathbf{w}^T \mathbf{x} \quad \text{(linear unit)}$$

$$\Delta \mathbf{w} \stackrel{(20)}{=} \eta(y\mathbf{x} - y^2\mathbf{w}) \quad \text{(Oja's rule)}$$

Linear Stability Analysis

$$
\begin{aligned}
y &\overset{(4)}{=} \mathbf{w}^T\mathbf{x} \quad \text{(linear unit)} \\
\Delta\mathbf{w} &\overset{(20)}{=} \eta(y\mathbf{x} - y^2\mathbf{w}) \quad \text{(Oja's rule)}
\end{aligned}
$$

---

1. What is the mean dynamics of the weight vectors?

$$\frac{1}{\eta}\langle\Delta\mathbf{w}\rangle_\mu \quad = \quad ?$$

2. What are the stationary weight vectors of the dynamics?

$$\mathbf{0} \overset{!}{=} \langle\Delta\mathbf{w}\rangle_\mu \quad \Longleftrightarrow \quad \mathbf{w} = ?$$

3. Which weight vectors are stable?

$$\mathbf{w} = ? \text{ stable} \quad \Longleftrightarrow \quad ?$$

What is the mean dynamics of the weight vectors?

$$y \overset{(4)}{=} \mathbf{w}^T \mathbf{x} \quad \text{(linear unit)}$$

$$(\Delta \mathbf{w})^\mu \overset{(20)}{=} \eta(y^\mu \mathbf{x}^\mu - (y^\mu)^2 \mathbf{w}) \quad \text{(Oja's rule)}$$

What is the mean dynamics of the weight vectors?

$$y \overset{(4)}{=} \mathbf{w}^T \mathbf{x} \quad \text{(linear unit)}$$

$$(\Delta \mathbf{w})^\mu \overset{(20)}{=} \eta(y^\mu \mathbf{x}^\mu - (y^\mu)^2 \mathbf{w}) \quad \text{(Oja's rule)}$$

In order to proceed analytically we have to average over all training patterns μ and get

$$\langle (\Delta \mathbf{w})^\mu \rangle_\mu = \frac{1}{M} \sum_{\mu=1}^{M} (\Delta \mathbf{w})^\mu \tag{23}$$

$$\overset{(20)}{=} \frac{1}{M} \sum_{\mu=1}^{M} \eta(y^\mu \mathbf{x}^\mu - (y^\mu)^2 \mathbf{w}) \tag{24}$$

$$\Longleftrightarrow \quad \langle \Delta \mathbf{w} \rangle_\mu = \eta \langle y \mathbf{x} - y^2 \mathbf{w} \rangle_\mu \tag{25}$$

For small learning rates η this is a good approximation to the real training procedure (apart from a factor of M).

What is the mean dynamics of the weight vectors?

$$y \overset{(4)}{=} \mathbf{w}^T \mathbf{x}\,, \quad \text{(linear unit)}$$

$$\langle \Delta \mathbf{w} \rangle_\mu \overset{(25)}{=} \eta \langle y\mathbf{x} - y^2 \mathbf{w} \rangle_\mu\,. \quad \text{(Oja's rule, averaged)}$$

What is the mean dynamics of the weight vectors?

$$y \stackrel{(4)}{=} \mathbf{w}^T\mathbf{x}\,, \quad \text{(linear unit)}$$

$$\langle\Delta\mathbf{w}\rangle_\mu \stackrel{(25)}{=} \eta\langle y\mathbf{x} - y^2\mathbf{w}\rangle_\mu\,. \quad \text{(Oja's rule, averaged)}$$

$$\frac{1}{\eta}\langle\Delta\mathbf{w}\rangle_\mu \stackrel{(25)}{=} \langle y\mathbf{x} - y^2\mathbf{w}\rangle_\mu \tag{26}$$

$$\stackrel{(4)}{=} \left\langle(\mathbf{w}^T\mathbf{x})\mathbf{x} - (\mathbf{w}^T\mathbf{x})(\mathbf{w}^T\mathbf{x})\mathbf{w}\right\rangle_\mu \tag{27}$$

$$= \left\langle\mathbf{x}(\mathbf{x}^T\mathbf{w}) - (\mathbf{w}^T\mathbf{x})(\mathbf{x}^T\mathbf{w})\mathbf{w}\right\rangle_\mu \tag{28}$$

$$= \left\langle(\mathbf{x}\mathbf{x}^T)\mathbf{w} - (\mathbf{w}^T(\mathbf{x}\mathbf{x}^T)\mathbf{w})\mathbf{w}\right\rangle_\mu \tag{29}$$

$$= \langle\mathbf{x}\mathbf{x}^T\rangle_\mu\mathbf{w} - (\mathbf{w}^T\langle\mathbf{x}\mathbf{x}^T\rangle_\mu\mathbf{w})\mathbf{w} \tag{30}$$

$$\left[\mathbf{C} := \langle\mathbf{x}\mathbf{x}^T\rangle_\mu = \mathbf{C}^T\right] \tag{31}$$

$$\stackrel{(31)}{=} \mathbf{C}\mathbf{w} - \left(\mathbf{w}^T\mathbf{C}\mathbf{w}\right)\mathbf{w} \tag{32}$$

Interim Summary

$$y \overset{(4)}{=} \mathbf{w}^T\mathbf{x} \quad \text{(linear unit)}$$

$$\Delta\mathbf{w} \overset{(20)}{=} \eta(y\mathbf{x} - y^2\mathbf{w}) \quad \text{(Oja's rule)}$$

---

1. What is the mean dynamics of the weight vectors?

$$\frac{1}{\eta}\langle\Delta\mathbf{w}\rangle_\mu \overset{(32)}{=} \mathbf{C}\mathbf{w} - \left(\mathbf{w}^T\mathbf{C}\mathbf{w}\right)\mathbf{w}$$

$$\text{with} \quad \mathbf{C} \overset{(31)}{:=} \langle\mathbf{x}\mathbf{x}^T\rangle_\mu$$

2. What are the stationary weight vectors of the dynamics?

$$\mathbf{0} \overset{!}{=} \langle\Delta\mathbf{w}\rangle_\mu \iff \mathbf{w} = ?$$

3. Which weight vectors are stable?

$$\mathbf{w} = ? \text{ stable} \iff ?$$

What are the stationary weight vectors?

$$\mathbf{C} \stackrel{(31)}{=} \langle \mathbf{x}\mathbf{x}^T \rangle_\mu = \mathbf{C}^T \,, \quad \text{(symmetric matrix } \mathbf{C}\text{)}$$

$$\frac{1}{\eta}\langle \Delta \mathbf{w} \rangle_\mu \stackrel{(32)}{=} \mathbf{C}\mathbf{w} - \left(\mathbf{w}^T \mathbf{C}\mathbf{w}\right) \mathbf{w} \,. \quad \text{(weight dynamics)}$$

What are the stationary weight vectors?

$$\mathbf{C} \overset{(31)}{=} \langle \mathbf{x}\mathbf{x}^T \rangle_\mu = \mathbf{C}^T \,, \quad \text{(symmetric matrix } \mathbf{C}\text{)}$$

$$\frac{1}{\eta}\langle \Delta\mathbf{w} \rangle_\mu \overset{(32)}{=} \mathbf{C}\mathbf{w} - \left(\mathbf{w}^T\mathbf{C}\mathbf{w}\right)\mathbf{w} \,. \quad \text{(weight dynamics)}$$

For stationary weight vectors we have:

$$\mathbf{0} \overset{!}{=} \langle \Delta\mathbf{w} \rangle_\mu \tag{33}$$

$$\overset{(32)}{\Longleftrightarrow} \quad \mathbf{C}\mathbf{w} = \left(\mathbf{w}^T\mathbf{C}\mathbf{w}\right)\mathbf{w} \tag{34}$$

$$\left[\lambda := \left(\mathbf{w}^T\mathbf{C}\mathbf{w}\right)\right] \tag{35}$$

$$\overset{(35)}{\Longleftrightarrow} \quad \mathbf{C}\mathbf{w} = \lambda\mathbf{w} \tag{36}$$

What are the stationary weight vectors?

$$\begin{aligned} \mathbf{C} &\stackrel{(31)}{=} \langle \mathbf{x}\mathbf{x}^T \rangle_\mu = \mathbf{C}^T \,, \quad \text{(symmetric matrix } \mathbf{C}\text{)} \\ \frac{1}{\eta} \langle \Delta \mathbf{w} \rangle_\mu &\stackrel{(32)}{=} \mathbf{C}\mathbf{w} - \left(\mathbf{w}^T \mathbf{C} \mathbf{w} \right) \mathbf{w} \,. \quad \text{(weight dynamics)} \end{aligned}$$

For stationary weight vectors we have:

$$\mathbf{0} \stackrel{!}{=} \langle \Delta \mathbf{w} \rangle_\mu \tag{33}$$

$$\stackrel{(32)}{\Longleftrightarrow} \quad \mathbf{C}\mathbf{w} = \left(\mathbf{w}^T \mathbf{C} \mathbf{w} \right) \mathbf{w} \tag{34}$$

$$\left[\lambda := \left(\mathbf{w}^T \mathbf{C} \mathbf{w} \right) \right] \tag{35}$$

$$\stackrel{(35)}{\Longleftrightarrow} \quad \mathbf{C}\mathbf{w} = \lambda \mathbf{w} \tag{36}$$

Stationary weight vectors **w** are eigenvectors of matrix **C**.

What are the stationary weight vectors?

$$\mathbf{C} \stackrel{(31)}{=} \langle \mathbf{x}\mathbf{x}^T \rangle_\mu = \mathbf{C}^T \,, \quad \text{(symmetric matrix } \mathbf{C}\text{)}$$

$$\frac{1}{\eta}\langle \Delta\mathbf{w} \rangle_\mu \stackrel{(32)}{=} \mathbf{C}\mathbf{w} - \left(\mathbf{w}^T\mathbf{C}\mathbf{w}\right)\mathbf{w}\,. \quad \text{(weight dynamics)}$$

For stationary weight vectors we have:

$$\mathbf{0} \stackrel{!}{=} \langle \Delta\mathbf{w} \rangle_\mu \tag{33}$$

$$\stackrel{(32)}{\iff} \quad \mathbf{C}\mathbf{w} = \left(\mathbf{w}^T\mathbf{C}\mathbf{w}\right)\mathbf{w} \tag{34}$$

$$\left[\lambda := \left(\mathbf{w}^T\mathbf{C}\mathbf{w}\right)\right] \tag{35}$$

$$\stackrel{(35)}{\iff} \quad \mathbf{C}\mathbf{w} = \lambda\mathbf{w} \tag{36}$$

Stationary weight vectors **w** are eigenvectors of matrix **C**.

$$\lambda \stackrel{(35)}{=} \mathbf{w}^T\mathbf{C}\mathbf{w} \stackrel{(36)}{=} \mathbf{w}^T\lambda\mathbf{w} = \lambda\mathbf{w}^T\mathbf{w} = \lambda\|\, \mathbf{w} \,\|^2 \tag{37}$$

$$\iff \quad 1 = \|\, \mathbf{w} \,\|^2 \tag{38}$$

Stationary weight vectors **w** are normalized to 1.

Interim Summary

$$y \overset{(4)}{=} \mathbf{w}^T \mathbf{x} \quad \text{(linear unit)}$$
$$\Delta \mathbf{w} \overset{(20)}{=} \eta (y\mathbf{x} - y^2 \mathbf{w}) \quad \text{(Oja's rule)}$$

---

1. What is the mean dynamics of the weight vectors?

$$\frac{1}{\eta} \langle \Delta \mathbf{w} \rangle_\mu \overset{(32)}{=} \mathbf{C}\mathbf{w} - \left(\mathbf{w}^T \mathbf{C} \mathbf{w} \right) \mathbf{w}$$
$$\text{with} \quad \mathbf{C} \overset{(31)}{:=} \langle \mathbf{x}\mathbf{x}^T \rangle_\mu$$

2. What are the stationary weight vectors of the dynamics?

$$\mathbf{0} \overset{!}{=} \langle \Delta \mathbf{w} \rangle_\mu \iff \mathbf{w} = \mathbf{c}^\alpha$$
$$\text{with } \mathbf{C}\mathbf{c}^\alpha \overset{(36)}{=} \lambda_\alpha \mathbf{c}^\alpha \quad \text{(eigenvectors of } \mathbf{C}\text{)}$$
$$\wedge \quad 1 \overset{(38)}{=} \| \mathbf{c}^\alpha \|^2 \quad \text{(with norm 1)}$$

3. Which weight vectors are stable?

$$\mathbf{w} = ? \text{ stable} \iff ?$$

Reminder: Eigenvalues and Eigenvectors

$$\mathbf{Aa} = \lambda \mathbf{a} \quad \text{(eigenvalue equation)} \tag{39}$$

Solutions of the eigenvalue equation for a given quadratic $N \times N$-matrix $\mathbf{A}$ are called eigenvectors $\mathbf{a}$ and eigenvalues λ.

For symmetric $\mathbf{A}$, i.e. $\mathbf{A}^T = \mathbf{A}$, eigenvalues λ are real, and eigenvectors to different eigenvalues are orthogonal, i.e. $\lambda_\alpha \neq \lambda_\beta \Rightarrow \mathbf{a}^\alpha \perp \mathbf{a}^\beta$.

A symmetric matrix $\mathbf{A}$ always has a complete set of orthonormal eigenvectors $\mathbf{a}^\alpha, \alpha = 1, ..., N$ (orthonormal basis), i.e.

$$\mathbf{Aa}^\alpha = \lambda_\alpha \mathbf{a}^\alpha \, , \quad \text{(right-eigenvectors)} \tag{40}$$

$$\iff \quad \mathbf{a}^{\alpha\, T}\mathbf{A} = (\mathbf{Aa}^\alpha)^T = \mathbf{a}^{\alpha\, T}\lambda_\alpha \, , \quad \text{(left-eigenvectors)} \tag{41}$$

$$\| \mathbf{a}^\alpha \| = \sqrt{\mathbf{a}^{\alpha\, T}\mathbf{a}^\alpha} = 1 \, , \quad \text{(with norm 1)} \tag{42}$$

$$\mathbf{a}^{\alpha\, T}\mathbf{a}^\beta = 0 \quad \forall \alpha \neq \beta \, . \quad \text{(orthogonal)} \tag{43}$$

$$\forall \mathbf{v} \quad \mathbf{v} = \sum_{\alpha=1}^{N} v'_\alpha \mathbf{a}^\alpha \quad \text{mit} \quad v'_\alpha = \mathbf{a}^{\alpha\, T}\mathbf{v} \quad \text{(complete)} \tag{44}$$

$$\Longrightarrow \quad \forall \mathbf{v} \quad \mathbf{v} = \sum_{\alpha=1}^{N} \mathbf{a}^\alpha \mathbf{a}^{\alpha\, T}\mathbf{v} \quad \Leftrightarrow \quad \mathbf{1} = \sum_{\alpha=1}^{N} \mathbf{a}^\alpha \mathbf{a}^{\alpha\, T} \tag{45}$$

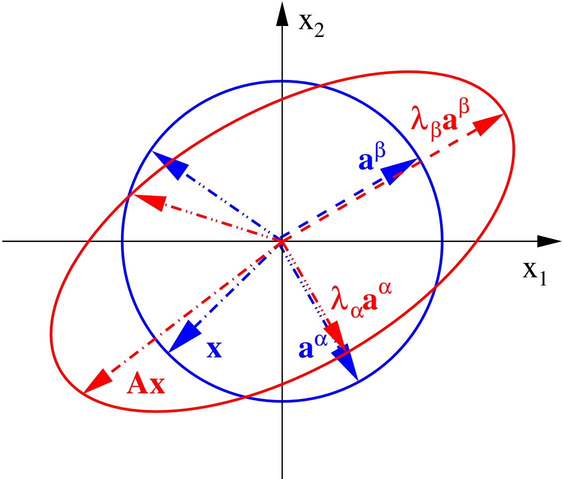
$$\| \mathbf{v} \|^2 = \sum_i v_i^2 \overset{(42,\,43,\,44)}{=} \sum_\alpha {v'_\alpha}^2 \tag{46}$$

Reminder: Eigenvalues and Eigenvectors

$$\mathbf{A}^T = \mathbf{A}\,, \quad \text{(symmetric)} \tag{47}$$

$$\mathbf{A}\mathbf{a}^\alpha = \lambda_\alpha \mathbf{a}^\alpha\,, \quad \text{(eigenvectors)} \tag{48}$$

$$\mathbf{a}^{\alpha\,T}\mathbf{a}^\beta = \delta_{\alpha\beta} \quad \forall \alpha, \beta\,. \quad \text{(orthonormal)} \tag{49}$$



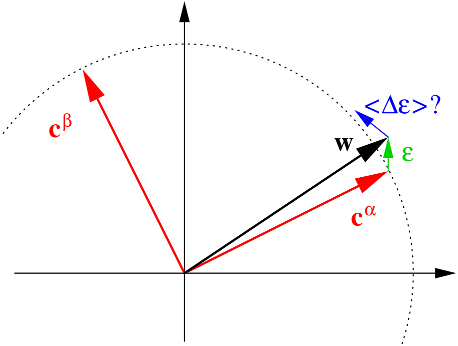
Which weight vectors are stable?

$$\frac{1}{\eta}\langle\Delta\mathbf{w}\rangle_{\mu} \overset{(32)}{=} \mathbf{C}\mathbf{w} - \left(\mathbf{w}^{T}\mathbf{C}\mathbf{w}\right)\mathbf{w} \tag{50}$$

Which weight vectors are stable?

$$\frac{1}{\eta}\langle\Delta \mathbf{w}\rangle_\mu \stackrel{(32)}{=} \mathbf{C}\mathbf{w} - \left(\mathbf{w}^T\mathbf{C}\mathbf{w}\right)\mathbf{w} \qquad (50)$$

$$\text{Ansatz:} \quad \mathbf{w} := \mathbf{c}^\alpha + \epsilon \quad \text{(with small } \epsilon) \qquad (51)$$



How does ϵ develop under the dynamics?

Which weight vectors are stable?

$$\frac{1}{\eta}\langle\Delta\mathbf{w}\rangle_\mu \overset{(32)}{=} \mathbf{C}\mathbf{w} - \left(\mathbf{w}^T\mathbf{C}\mathbf{w}\right)\mathbf{w}$$

$$\mathbf{w} \overset{(51)}{:=} \mathbf{c}^\alpha + \epsilon \quad \text{(with small } \epsilon\text{)}$$

$$\begin{aligned}
\frac{1}{\eta}\langle\Delta\epsilon\rangle_\mu &\overset{(51)}{=} \frac{1}{\eta}\langle\Delta\mathbf{w}\rangle_\mu && (52)\\
&\overset{(32)}{=} \mathbf{C}\mathbf{w} - \left(\mathbf{w}^T\mathbf{C}\mathbf{w}\right)\mathbf{w} && (53)\\
&\overset{(51)}{=} \mathbf{C}(\mathbf{c}^\alpha + \epsilon) - \left((\mathbf{c}^\alpha + \epsilon)^T\mathbf{C}(\mathbf{c}^\alpha + \epsilon)\right)(\mathbf{c}^\alpha + \epsilon) && (54)\\
&\approx \mathbf{C}\mathbf{c}^\alpha + \mathbf{C}\epsilon - (\mathbf{c}^{\alpha T}\mathbf{C}\mathbf{c}^\alpha)\mathbf{c}^\alpha - (\mathbf{c}^{\alpha T}\mathbf{C}\epsilon)\mathbf{c}^\alpha\\
&\quad - (\epsilon^T\mathbf{C}\mathbf{c}^\alpha)\mathbf{c}^\alpha - (\mathbf{c}^{\alpha T}\mathbf{C}\mathbf{c}^\alpha)\epsilon && (55)\\
&\overset{(36,41)}{=} \lambda^\alpha\mathbf{c}^\alpha + \mathbf{C}\epsilon - (\mathbf{c}^{\alpha T}\lambda^\alpha\mathbf{c}^\alpha)\mathbf{c}^\alpha - (\mathbf{c}^{\alpha T}\lambda^\alpha\epsilon)\mathbf{c}^\alpha\\
&\quad - (\epsilon^T\lambda^\alpha\mathbf{c}^\alpha)\mathbf{c}^\alpha - (\mathbf{c}^{\alpha T}\lambda^\alpha\mathbf{c}^\alpha)\epsilon && (56)\\
&\overset{(38)}{=} \lambda^\alpha\mathbf{c}^\alpha + \mathbf{C}\epsilon - \lambda^\alpha\mathbf{c}^\alpha - \lambda^\alpha(\mathbf{c}^{\alpha T}\epsilon)\mathbf{c}^\alpha\\
&\quad - \lambda^\alpha(\epsilon^T\mathbf{c}^\alpha)\mathbf{c}^\alpha - \lambda^\alpha\epsilon && (57)\\
&= \mathbf{C}\epsilon - 2\lambda^\alpha(\mathbf{c}^{\alpha T}\epsilon)\mathbf{c}^\alpha - \lambda^\alpha\epsilon\,. && (58)
\end{aligned}$$

Which weight vectors are stable?

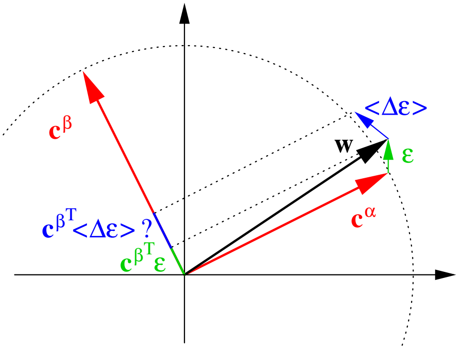
$$\mathbf{w} \overset{(51)}{:=} \mathbf{c}^{\alpha} + \epsilon\,,$$
$$\frac{1}{\eta}\langle \Delta\epsilon \rangle_{\mu} \overset{(58)}{\approx} \mathbf{C}\epsilon - 2\lambda^{\alpha}(\mathbf{c}^{\alpha\,T}\epsilon)\mathbf{c}^{\alpha} - \lambda^{\alpha}\epsilon\,.$$

Which weight vectors are stable?

$$\mathbf{w} \stackrel{(51)}{:=} \mathbf{c}^{\alpha} + \epsilon\,,$$

$$\frac{1}{\eta}\langle \Delta\epsilon \rangle_{\mu} \stackrel{(58)}{\approx} \mathbf{C}\epsilon - 2\lambda^{\alpha}(\mathbf{c}^{\alpha\,T}\epsilon)\mathbf{c}^{\alpha} - \lambda^{\alpha}\epsilon\,.$$

For simplicity consider the change of the perturbation along the eigenvector $\mathbf{c}^{\beta}$.



Which weight vectors are stable?

$$\begin{aligned}
\mathbf{w} &\overset{(51)}{:=} \mathbf{c}^\alpha + \epsilon\,, \\
\frac{1}{\eta}\langle\Delta\epsilon\rangle_\mu &\overset{(58)}{\approx} \mathbf{C}\epsilon - 2\lambda^\alpha(\mathbf{c}^{\alpha\,T}\epsilon)\mathbf{c}^\alpha - \lambda^\alpha\epsilon\,.
\end{aligned}$$

For simplicity consider the change of the perturbation along the eigenvector $\mathbf{c}^\beta$.

$$\begin{aligned}
\frac{1}{\eta}\mathbf{c}^{\beta\,T}\langle\Delta\epsilon\rangle_\mu &\overset{(58)}{\approx} \mathbf{c}^{\beta\,T}\mathbf{C}\epsilon - 2\lambda_\alpha(\mathbf{c}^{\alpha\,T}\epsilon)(\mathbf{c}^{\beta\,T}\mathbf{c}^\alpha) - \lambda_\alpha(\mathbf{c}^{\beta\,T}\epsilon) && (59) \\
&\overset{(36,41)}{=} \lambda_\beta(\mathbf{c}^{\beta\,T}\epsilon) - 2\lambda_\alpha(\mathbf{c}^{\alpha\,T}\epsilon)(\mathbf{c}^{\beta\,T}\mathbf{c}^\alpha) - \lambda_\alpha(\mathbf{c}^{\beta\,T}\epsilon) && (60) \\
&\overset{(42,43)}{=} \lambda_\beta(\mathbf{c}^{\beta\,T}\epsilon) - 2\lambda_\alpha(\mathbf{c}^{\alpha\,T}\epsilon)\delta_{\beta\alpha} - \lambda_\alpha(\mathbf{c}^{\beta\,T}\epsilon) && (61) \\
&= \lambda_\beta(\mathbf{c}^{\beta\,T}\epsilon) - 2\lambda_\alpha(\mathbf{c}^{\beta\,T}\epsilon)\delta_{\beta\alpha} - \lambda_\alpha(\mathbf{c}^{\beta\,T}\epsilon) && (62) \\
&= (\lambda_\beta - 2\lambda_\alpha\delta_{\beta\alpha} - \lambda_\alpha)\,(\mathbf{c}^{\beta\,T}\epsilon) && (63) \\
&= \begin{cases} (-2\lambda_\beta)(\mathbf{c}^{\beta\,T}\epsilon) & \text{if } \beta = \alpha \\ (\lambda_\beta - \lambda_\alpha)(\mathbf{c}^{\beta\,T}\epsilon) & \text{if } \beta \neq \alpha \end{cases} && (64)
\end{aligned}$$

Which weight vectors are stable?

$$\mathbf{w} \overset{(51)}{:=} \mathbf{c}^{\alpha} + \epsilon \,,$$

$$\frac{1}{\eta}{\mathbf{c}^{\beta}}^{T}\langle\Delta\epsilon\rangle_{\mu} \overset{(64)}{\approx} \begin{cases} (-2\lambda_{\beta})({\mathbf{c}^{\beta}}^{T}\epsilon) & \text{if } \beta = \alpha \\ (\lambda_{\beta} - \lambda_{\alpha})({\mathbf{c}^{\beta}}^{T}\epsilon) & \text{if } \beta \neq \alpha \end{cases} \,.$$

Which weight vectors are stable?

$$\mathbf{w} \overset{(51)}{:=} \mathbf{c}^\alpha + \epsilon\,,$$

$$\frac{1}{\eta}\mathbf{c}^{\beta\,T}\langle\Delta\epsilon\rangle_\mu \overset{(64)}{\approx} \begin{cases} (-2\lambda_\beta)(\mathbf{c}^{\beta\,T}\epsilon) & \text{if } \beta = \alpha \\ (\lambda_\beta - \lambda_\alpha)(\mathbf{c}^{\beta\,T}\epsilon) & \text{if } \beta \neq \alpha \end{cases} .$$

With

$$\mathbf{c}^{\beta\,T}\langle\Delta\epsilon\rangle_\mu = \langle\mathbf{c}^{\beta\,T}\Delta\epsilon\rangle_\mu \tag{65}$$

$$= \langle\mathbf{c}^{\beta\,T}(\epsilon^{n+1} - \epsilon^{n})\rangle_\mu \tag{66}$$

$$= \langle(\mathbf{c}^{\beta\,T}\epsilon)^{n+1} - (\mathbf{c}^{\beta\,T}\epsilon)^{n}\rangle_\mu \tag{67}$$

$$= \langle\Delta(\mathbf{c}^{\beta\,T}\epsilon)\rangle_\mu \tag{68}$$

Which weight vectors are stable?

$$\mathbf{w} \overset{(51)}{:=} \mathbf{c}^\alpha + \epsilon\,,$$

$$\frac{1}{\eta}\mathbf{c}^{\beta\,T}\langle\Delta\epsilon\rangle_\mu \overset{(64)}{\approx} \begin{cases} (-2\lambda_\beta)(\mathbf{c}^{\beta\,T}\epsilon) & \text{if } \beta = \alpha \\ (\lambda_\beta - \lambda_\alpha)(\mathbf{c}^{\beta\,T}\epsilon) & \text{if } \beta \neq \alpha \end{cases} .$$

With

$$\mathbf{c}^{\beta\,T}\langle\Delta\epsilon\rangle_\mu = \langle\mathbf{c}^{\beta\,T}\Delta\epsilon\rangle_\mu \quad (65)$$

$$= \langle\mathbf{c}^{\beta\,T}(\epsilon^{n+1} - \epsilon^n)\rangle_\mu \quad (66)$$

$$= \langle(\mathbf{c}^{\beta\,T}\epsilon)^{n+1} - (\mathbf{c}^{\beta\,T}\epsilon)^n\rangle_\mu \quad (67)$$

$$= \langle\Delta(\mathbf{c}^{\beta\,T}\epsilon)\rangle_\mu \quad (68)$$

$$s_\beta := \mathbf{c}^{\beta\,T}\epsilon \quad \text{(perturbation along } \mathbf{c}^\beta) \quad (69)$$

Which weight vectors are stable?

$$\mathbf{w} \overset{(51)}{:=} \mathbf{c}^{\alpha} + \epsilon\,,$$

$$\frac{1}{\eta}\mathbf{c}^{\beta^T}\langle\Delta\epsilon\rangle_{\mu} \overset{(64)}{\approx} \begin{cases} (-2\lambda_{\beta})(\mathbf{c}^{\beta^T}\epsilon) & \text{if } \beta = \alpha \\ (\lambda_{\beta} - \lambda_{\alpha})(\mathbf{c}^{\beta^T}\epsilon) & \text{if } \beta \neq \alpha \end{cases}\,.$$

With

$$\mathbf{c}^{\beta^T}\langle\Delta\epsilon\rangle_{\mu} = \langle\mathbf{c}^{\beta^T}\Delta\epsilon\rangle_{\mu} \tag{65}$$

$$= \langle\mathbf{c}^{\beta^T}(\epsilon^{n+1} - \epsilon^{n})\rangle_{\mu} \tag{66}$$

$$= \langle(\mathbf{c}^{\beta^T}\epsilon)^{n+1} - (\mathbf{c}^{\beta^T}\epsilon)^{n}\rangle_{\mu} \tag{67}$$

$$= \langle\Delta(\mathbf{c}^{\beta^T}\epsilon)\rangle_{\mu} \tag{68}$$

$$s_{\beta} := \mathbf{c}^{\beta^T}\epsilon \quad \text{(perturbation along } \mathbf{c}^{\beta}\text{)} \tag{69}$$

$$\kappa_{\alpha\beta} := \begin{cases} \eta(-2\lambda_{\beta}) & \text{if } \beta = \alpha \\ \eta(\lambda_{\beta} - \lambda_{\alpha}) & \text{if } \beta \neq \alpha \end{cases} \tag{70}$$

Which weight vectors are stable?

$$\begin{array}{rcl} \mathbf{w} & \stackrel{(51)}{:=} & \mathbf{c}^\alpha + \epsilon \,, \\ \frac{1}{\eta}{\mathbf{c}^\beta}^T \langle \Delta\epsilon \rangle_\mu & \stackrel{(64)}{\approx} & \left\{ \begin{array}{rl} (-2\lambda_\beta)({\mathbf{c}^\beta}^T\epsilon) & \text{if } \beta = \alpha \\ (\lambda_\beta - \lambda_\alpha)({\mathbf{c}^\beta}^T\epsilon) & \text{if } \beta \neq \alpha \end{array} \right. . \end{array}$$

With

$$\begin{aligned} {\mathbf{c}^\beta}^T \langle \Delta\epsilon \rangle_\mu &= \langle {\mathbf{c}^\beta}^T \Delta\epsilon \rangle_\mu && (65) \\ &= \langle {\mathbf{c}^\beta}^T (\epsilon^{n+1} - \epsilon^n) \rangle_\mu && (66) \\ &= \langle ({\mathbf{c}^\beta}^T\epsilon)^{n+1} - ({\mathbf{c}^\beta}^T\epsilon)^n \rangle_\mu && (67) \\ &= \langle \Delta({\mathbf{c}^\beta}^T\epsilon) \rangle_\mu && (68) \\ s_\beta &:= {\mathbf{c}^\beta}^T\epsilon \quad \text{(perturbation along } \mathbf{c}^\beta) && (69) \\ \kappa_{\alpha\beta} &:= \left\{ \begin{array}{rl} \eta(-2\lambda_\beta) & \text{if } \beta = \alpha \\ \eta(\lambda_\beta - \lambda_\alpha) & \text{if } \beta \neq \alpha \end{array} \right. && (70) \end{aligned}$$

we get

$$(64) \quad \stackrel{(68,\,69,\,70)}{\Longleftrightarrow} \quad \langle \Delta s_\beta \rangle_\mu \approx \kappa_{\alpha\beta} s_\beta \,. \qquad (71)$$

Which weight vectors are stable?

$$\mathbf{w} \quad \overset{(51)}{:=} \quad \mathbf{c}^{\alpha} + \epsilon \,,$$

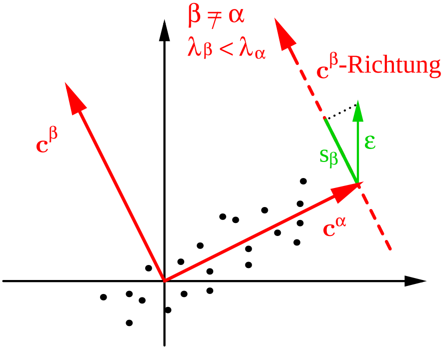
$$\langle \Delta s_{\beta} \rangle_{\mu} \overset{(71)}{\approx} \kappa_{\alpha\beta} s_{\beta} \quad \text{with} \quad \kappa_{\alpha\beta} \overset{(70)}{:=} \begin{cases} \eta(-2\lambda_{\beta}) & \text{if } \beta = \alpha \\ \eta(\lambda_{\beta} - \lambda_{\alpha}) & \text{if } \beta \neq \alpha \end{cases} \,.$$

Which weight vectors are stable?

$$\mathbf{w} \overset{(51)}{:=} \mathbf{c}^\alpha + \epsilon\,,$$

$$\langle \Delta s_\beta \rangle_\mu \overset{(71)}{\approx} \kappa_{\alpha\beta} s_\beta \quad \text{with} \quad \kappa_{\alpha\beta} \overset{(70)}{:=} \begin{cases} \eta(-2\lambda_\beta) & \text{if } \beta = \alpha \\ \eta(\lambda_\beta - \lambda_\alpha) & \text{if } \beta \neq \alpha \end{cases} .$$

Case 1: $\beta \neq \alpha, \lambda_\beta < \lambda_\alpha$

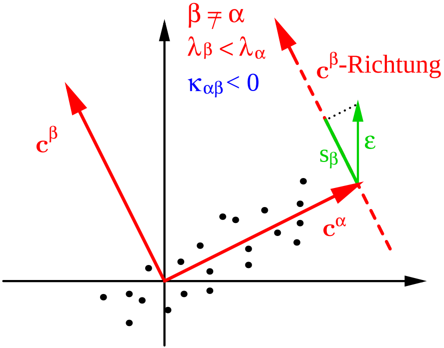


Which weight vectors are stable?

$$\mathbf{w} \overset{(51)}{:=} \mathbf{c}^\alpha + \epsilon\,,$$

$$\langle \Delta s_\beta \rangle_\mu \overset{(71)}{\approx} \kappa_{\alpha\beta} s_\beta \quad \text{with} \quad \kappa_{\alpha\beta} \overset{(70)}{:=} \begin{cases} \eta(-2\lambda_\beta) & \text{if } \beta = \alpha \\ \eta(\lambda_\beta - \lambda_\alpha) & \text{if } \beta \neq \alpha \end{cases} .$$

Case 1: $\beta \neq \alpha, \lambda_\beta < \lambda_\alpha \quad \Rightarrow \quad \kappa_{\alpha\beta} = \eta(\lambda_\beta - \lambda_\alpha) < 0$,

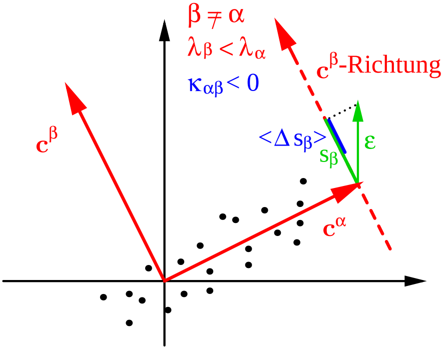


Which weight vectors are stable?

$$\mathbf{w} \stackrel{(51)}{:=} \mathbf{c}^\alpha + \epsilon\,,$$

$$\langle \Delta s_\beta \rangle_\mu \stackrel{(71)}{\approx} \kappa_{\alpha\beta} s_\beta \quad \text{with} \quad \kappa_{\alpha\beta} \stackrel{(70)}{:=} \begin{cases} \eta(-2\lambda_\beta) & \text{if } \beta = \alpha \\ \eta(\lambda_\beta - \lambda_\alpha) & \text{if } \beta \neq \alpha \end{cases} .$$

Case 1: $\beta \neq \alpha, \lambda_\beta < \lambda_\alpha \quad \Rightarrow \quad \kappa_{\alpha\beta} = \eta(\lambda_\beta - \lambda_\alpha) < 0$,
The perturbation along $\mathbf{c}^\beta$ decays.



Which weight vectors are stable?

$$\mathbf{w} \overset{(51)}{:=} \mathbf{c}^{\alpha} + \epsilon \,,$$

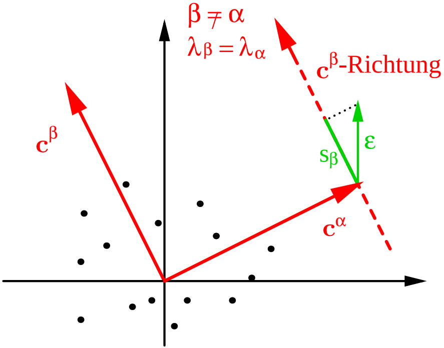
$$\langle \Delta s_{\beta} \rangle_{\mu} \overset{(71)}{\approx} \kappa_{\alpha\beta} s_{\beta} \quad \text{with} \quad \kappa_{\alpha\beta} \overset{(70)}{:=} \begin{cases} \eta(-2\lambda_{\beta}) & \text{if } \beta = \alpha \\ \eta(\lambda_{\beta} - \lambda_{\alpha}) & \text{if } \beta \neq \alpha \end{cases} \,.$$

Which weight vectors are stable?

$$\mathbf{w} \stackrel{(51)}{:=} \mathbf{c}^\alpha + \epsilon\,,$$

$$\langle \Delta s_\beta \rangle_\mu \stackrel{(71)}{\approx} \kappa_{\alpha\beta} s_\beta \quad \text{with} \quad \kappa_{\alpha\beta} \stackrel{(70)}{:=} \begin{cases} \eta(-2\lambda_\beta) & \text{if } \beta = \alpha \\ \eta(\lambda_\beta - \lambda_\alpha) & \text{if } \beta \neq \alpha \end{cases} .$$

Case 2: $\beta \neq \alpha, \lambda_\beta = \lambda_\alpha$

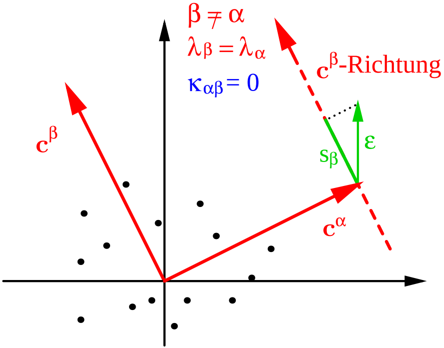


Which weight vectors are stable?

$$\mathbf{w} \overset{(51)}{:=} \mathbf{c}^{\alpha} + \epsilon\,,$$

$$\langle \Delta s_\beta \rangle_\mu \overset{(71)}{\approx} \kappa_{\alpha\beta} s_\beta \quad \text{with} \quad \kappa_{\alpha\beta} \overset{(70)}{:=} \begin{cases} \eta(-2\lambda_\beta) & \text{if } \beta = \alpha \\ \eta(\lambda_\beta - \lambda_\alpha) & \text{if } \beta \neq \alpha \end{cases} \,.$$

Case 2: $\beta \neq \alpha, \lambda_\beta = \lambda_\alpha \quad \Rightarrow \quad \kappa_{\alpha\beta} = \eta(\lambda_\beta - \lambda_\alpha) = 0,$

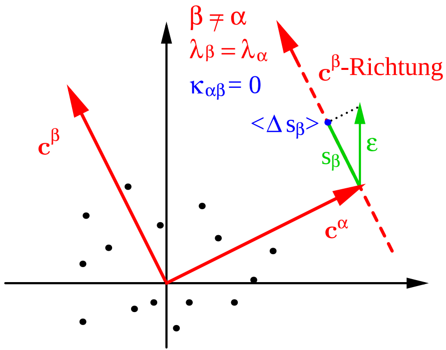


Which weight vectors are stable?

$$\mathbf{w} \overset{(51)}{:=} \mathbf{c}^{\alpha} + \epsilon\,,$$

$$\langle \Delta s_\beta \rangle_\mu \overset{(71)}{\approx} \kappa_{\alpha\beta} s_\beta \quad \text{with} \quad \kappa_{\alpha\beta} \overset{(70)}{:=} \begin{cases} \eta(-2\lambda_\beta) & \text{if } \beta = \alpha \\ \eta(\lambda_\beta - \lambda_\alpha) & \text{if } \beta \neq \alpha \end{cases} .$$

Case 2: $\beta \neq \alpha, \lambda_\beta = \lambda_\alpha \quad \Rightarrow \quad \kappa_{\alpha\beta} = \eta(\lambda_\beta - \lambda_\alpha) = 0$,
The perturbation along $\mathbf{c}^\beta$ persists.



Which weight vectors are stable?

$$\mathbf{w} \overset{(51)}{:=} \mathbf{c}^{\alpha} + \epsilon\,,$$

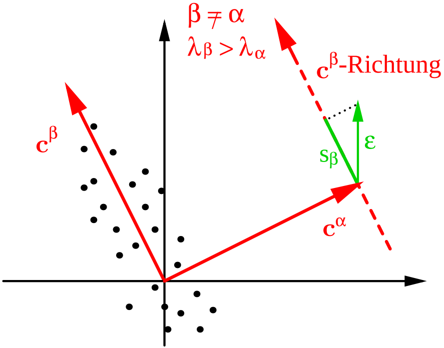
$$\langle \Delta s_\beta \rangle_\mu \overset{(71)}{\approx} \kappa_{\alpha\beta} s_\beta \quad \text{with} \quad \kappa_{\alpha\beta} \overset{(70)}{:=} \begin{cases} \eta(-2\lambda_\beta) & \text{if } \beta = \alpha \\ \eta(\lambda_\beta - \lambda_\alpha) & \text{if } \beta \neq \alpha \end{cases} \,.$$

Which weight vectors are stable?

$$\mathbf{w} \overset{(51)}{:=} \mathbf{c}^{\alpha} + \epsilon\,,$$

$$\langle \Delta s_\beta \rangle_\mu \overset{(71)}{\approx} \kappa_{\alpha\beta} s_\beta \quad \text{with} \quad \kappa_{\alpha\beta} \overset{(70)}{:=} \begin{cases} \eta(-2\lambda_\beta) & \text{if } \beta = \alpha \\ \eta(\lambda_\beta - \lambda_\alpha) & \text{if } \beta \neq \alpha \end{cases} \,.$$

Case 3: $\beta \neq \alpha, \lambda_\beta > \lambda_\alpha$

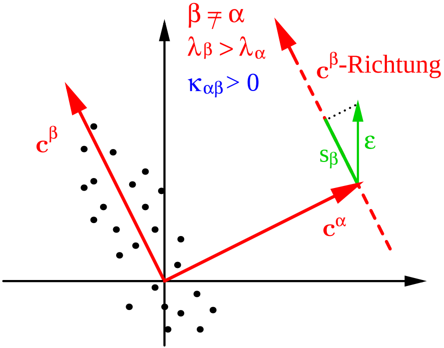


Which weight vectors are stable?

$$\mathbf{w} \overset{(51)}{:=} \mathbf{c}^\alpha + \epsilon\,,$$

$$\langle \Delta s_\beta \rangle_\mu \overset{(71)}{\approx} \kappa_{\alpha\beta} s_\beta \quad \text{with} \quad \kappa_{\alpha\beta} \overset{(70)}{:=} \begin{cases} \eta(-2\lambda_\beta) & \text{if } \beta = \alpha \\ \eta(\lambda_\beta - \lambda_\alpha) & \text{if } \beta \neq \alpha \end{cases} .$$

Case 3: $\beta \neq \alpha, \lambda_\beta > \lambda_\alpha \quad \Rightarrow \quad \kappa_{\alpha\beta} = \eta(\lambda_\beta - \lambda_\alpha) > 0$,

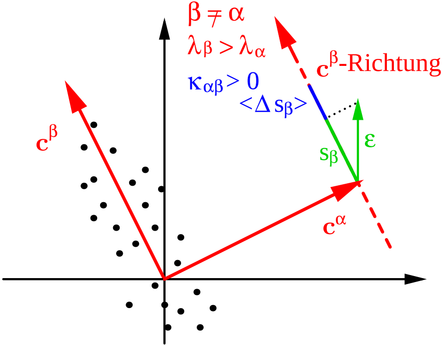


Which weight vectors are stable?

$$\mathbf{w} \overset{(51)}{:=} \mathbf{c}^\alpha + \epsilon\,,$$

$$\langle \Delta s_\beta \rangle_\mu \overset{(71)}{\approx} \kappa_{\alpha\beta} s_\beta \quad \text{with} \quad \kappa_{\alpha\beta} \overset{(70)}{:=} \begin{cases} \eta(-2\lambda_\beta) & \text{if } \beta = \alpha \\ \eta(\lambda_\beta - \lambda_\alpha) & \text{if } \beta \neq \alpha \end{cases} .$$

Case 3: $\beta \neq \alpha, \lambda_\beta > \lambda_\alpha \quad \Rightarrow \quad \kappa_{\alpha\beta} = \eta(\lambda_\beta - \lambda_\alpha) > 0$,
The perturbation along $\mathbf{c}^\beta$ grows.



Which weight vectors are stable?

$$\mathbf{w} \overset{(51)}{:=} \mathbf{c}^{\alpha} + \epsilon\,,$$

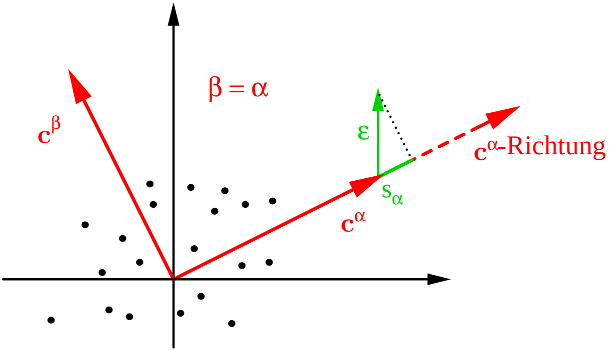
$$\langle \Delta s_\beta \rangle_\mu \overset{(71)}{\approx} \kappa_{\alpha\beta} s_\beta \quad \text{with} \quad \kappa_{\alpha\beta} \overset{(70)}{:=} \begin{cases} \eta(-2\lambda_\beta) & \text{if } \beta = \alpha \\ \eta(\lambda_\beta - \lambda_\alpha) & \text{if } \beta \neq \alpha \end{cases} .$$

Which weight vectors are stable?

$$\mathbf{w} \overset{(51)}{:=} \mathbf{c}^{\alpha} + \epsilon\,,$$

$$\langle \Delta s_\beta \rangle_\mu \overset{(71)}{\approx} \kappa_{\alpha\beta} s_\beta \quad \text{with} \quad \kappa_{\alpha\beta} \overset{(70)}{:=} \begin{cases} \eta(-2\lambda_\beta) & \text{if } \beta = \alpha \\ \eta(\lambda_\beta - \lambda_\alpha) & \text{if } \beta \neq \alpha \end{cases} .$$

Case 4: $\beta = \alpha$

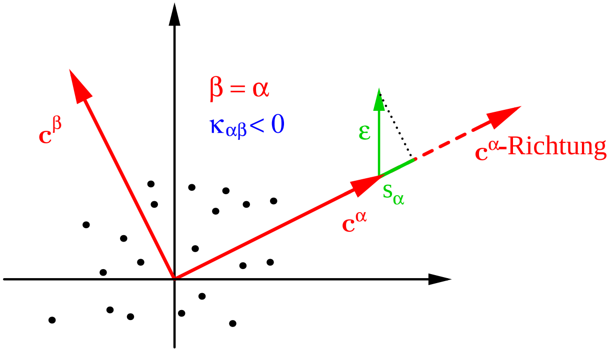


Which weight vectors are stable?

$$\mathbf{w} \overset{(51)}{:=} \mathbf{c}^\alpha + \epsilon\,,$$

$$\langle \Delta s_\beta \rangle_\mu \overset{(71)}{\approx} \kappa_{\alpha\beta} s_\beta \quad \text{with} \quad \kappa_{\alpha\beta} \overset{(70)}{:=} \begin{cases} \eta(-2\lambda_\beta) & \text{if } \beta = \alpha \\ \eta(\lambda_\beta - \lambda_\alpha) & \text{if } \beta \neq \alpha \end{cases} .$$

Case 4: $\beta = \alpha \quad \Rightarrow \quad \kappa_{\alpha\beta} = \eta(-2\lambda_\beta) < 0.$

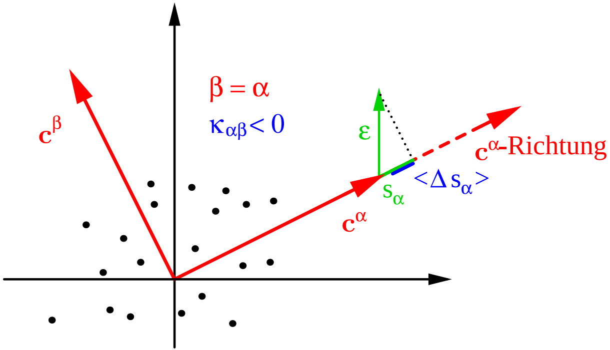


Which weight vectors are stable?

$$\mathbf{w} \stackrel{(51)}{:=} \mathbf{c}^\alpha + \epsilon\,,$$

$$\langle \Delta s_\beta \rangle_\mu \stackrel{(71)}{\approx} \kappa_{\alpha\beta} s_\beta \quad \text{with} \quad \kappa_{\alpha\beta} \stackrel{(70)}{:=} \begin{cases} \eta(-2\lambda_\beta) & \text{if } \beta = \alpha \\ \eta(\lambda_\beta - \lambda_\alpha) & \text{if } \beta \neq \alpha \end{cases} .$$

Case 4: $\beta = \alpha \quad \Rightarrow \quad \kappa_{\alpha\beta} = \eta(-2\lambda_\beta) < 0$.
The perturbation along $\mathbf{c}^\alpha$ always decays (if $\lambda_\beta > 0$).



Which weight vectors are stable?

$$\mathbf{w} \overset{(51)}{:=} \mathbf{c}^{\alpha} + \epsilon \,,$$

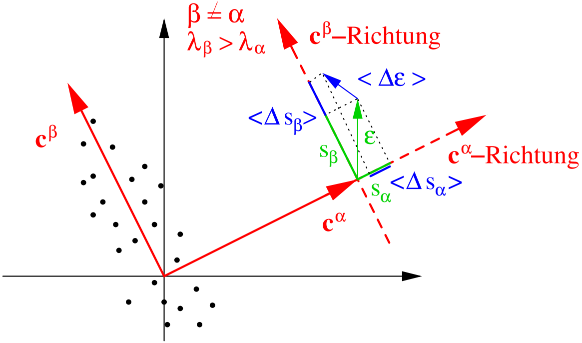
$$\langle \Delta s_{\beta} \rangle_{\mu} \overset{(71)}{\approx} \kappa_{\alpha\beta} s_{\beta} \quad \text{with} \quad \kappa_{\alpha\beta} \overset{(70)}{:=} \begin{cases} \eta(-2\lambda_{\beta}) & \text{if } \beta = \alpha \\ \eta(\lambda_{\beta} - \lambda_{\alpha}) & \text{if } \beta \neq \alpha \end{cases} \,.$$

Which weight vectors are stable?

$$\mathbf{w} \overset{(51)}{:=} \mathbf{c}^\alpha + \epsilon\,,$$

$$\langle \Delta s_\beta \rangle_\mu \overset{(71)}{\approx} \kappa_{\alpha\beta} s_\beta \quad \text{with} \quad \kappa_{\alpha\beta} \overset{(70)}{:=} \begin{cases} \eta(-2\lambda_\beta) & \text{if } \beta = \alpha \\ \eta(\lambda_\beta - \lambda_\alpha) & \text{if } \beta \neq \alpha \end{cases} .$$

Cases 3 and 4 combined.
The weight vectors turns away $\mathbf{c}^\alpha$ into the $\mathbf{c}^\beta$-direction.



Which weight vectors are stable?

$$\mathbf{w} \overset{(51)}{:=} \mathbf{c}^{\alpha} + \epsilon\,,$$

$$\langle \Delta s_{\beta} \rangle_{\mu} \overset{(71)}{\approx} \kappa_{\alpha\beta} s_{\beta} \quad \text{with} \quad \kappa_{\alpha\beta} \overset{(70)}{:=} \begin{cases} \eta(-2\lambda_{\beta}) & \text{if } \beta = \alpha \\ \eta(\lambda_{\beta} - \lambda_{\alpha}) & \text{if } \beta \neq \alpha \end{cases} .$$

Which weight vectors are stable?

$$\mathbf{w} \stackrel{(51)}{:=} \mathbf{c}^\alpha + \epsilon\,,$$

$$\langle \Delta s_\beta \rangle_\mu \stackrel{(71)}{\approx} \kappa_{\alpha\beta} s_\beta \quad \text{with} \quad \kappa_{\alpha\beta} \stackrel{(70)}{:=} \begin{cases} \eta(-2\lambda_\beta) & \text{if } \beta = \alpha \\ \eta(\lambda_\beta - \lambda_\alpha) & \text{if } \beta \neq \alpha \end{cases} .$$

Case 1: $\beta \neq \alpha, \lambda_\beta < \lambda_\alpha$, $\rightarrow$ The perturbation along $\mathbf{c}^\beta$ decays.

Case 2: $\beta \neq \alpha, \lambda_\beta = \lambda_\alpha$, $\rightarrow$ The perturbation along $\mathbf{c}^\beta$ persists.

Case 3: $\beta \neq \alpha, \lambda_\beta > \lambda_\alpha$, $\rightarrow$ The perturbation along $\mathbf{c}^\beta$ grows.

Case 4: $\beta = \alpha$, $\rightarrow$ The perturbation along $\mathbf{c}^\alpha$ always decays.

Which weight vectors are stable?

$$\mathbf{w} \overset{(51)}{:=} \mathbf{c}^\alpha + \epsilon\,,$$

$$\langle \Delta s_\beta \rangle_\mu \overset{(71)}{\approx} \kappa_{\alpha\beta} s_\beta \quad \text{with} \quad \kappa_{\alpha\beta} \overset{(70)}{:=} \begin{cases} \eta(-2\lambda_\beta) & \text{if } \beta = \alpha \\ \eta(\lambda_\beta - \lambda_\alpha) & \text{if } \beta \neq \alpha \end{cases} .$$

Case 1: $\beta \neq \alpha, \lambda_\beta < \lambda_\alpha$, $\rightarrow$ The perturbation along $\mathbf{c}^\beta$ decays.

Case 2: $\beta \neq \alpha, \lambda_\beta = \lambda_\alpha$, $\rightarrow$ The perturbation along $\mathbf{c}^\beta$ persists.

Case 3: $\beta \neq \alpha, \lambda_\beta > \lambda_\alpha$, $\rightarrow$ The perturbation along $\mathbf{c}^\beta$ grows.

Case 4: $\beta = \alpha$, $\rightarrow$ The perturbation along $\mathbf{c}^\alpha$ always decays.

The weight vector $\mathbf{w} = \mathbf{c}^\alpha$ is stable only if perturbations in all directions decay,

Which weight vectors are stable?

$$\mathbf{w} \stackrel{(51)}{:=} \mathbf{c}^\alpha + \epsilon\,,$$

$$\langle \Delta s_\beta \rangle_\mu \stackrel{(71)}{\approx} \kappa_{\alpha\beta} s_\beta \quad \text{with} \quad \kappa_{\alpha\beta} \stackrel{(70)}{:=} \begin{cases} \eta(-2\lambda_\beta) & \text{if } \beta = \alpha \\ \eta(\lambda_\beta - \lambda_\alpha) & \text{if } \beta \neq \alpha \end{cases} .$$

Case 1: $\beta \neq \alpha, \lambda_\beta < \lambda_\alpha$, $\rightarrow$ The perturbation along $\mathbf{c}^\beta$ decays.

Case 2: $\beta \neq \alpha, \lambda_\beta = \lambda_\alpha$, $\rightarrow$ The perturbation along $\mathbf{c}^\beta$ persists.

Case 3: $\beta \neq \alpha, \lambda_\beta > \lambda_\alpha$, $\rightarrow$ The perturbation along $\mathbf{c}^\beta$ grows.

Case 4: $\beta = \alpha$, $\rightarrow$ The perturbation along $\mathbf{c}^\alpha$ always decays.

The weight vector $\mathbf{w} = \mathbf{c}^\alpha$ is stable only if perturbations in all directions decay, i.e.

$$\mathbf{w} = \mathbf{c}^\alpha \text{ stable} \iff \lambda_\beta < \lambda_\alpha \quad \forall \beta \neq \alpha\,. \tag{72}$$

Which weight vectors are stable?

$$\mathbf{w} \overset{(51)}{:=} \mathbf{c}^\alpha + \epsilon\,,$$

$$\langle \Delta s_\beta \rangle_\mu \overset{(71)}{\approx} \kappa_{\alpha\beta} s_\beta \quad \text{with} \quad \kappa_{\alpha\beta} \overset{(70)}{:=} \begin{cases} \eta(-2\lambda_\beta) & \text{if } \beta = \alpha \\ \eta(\lambda_\beta - \lambda_\alpha) & \text{if } \beta \neq \alpha \end{cases} .$$

Case 1: $\beta \neq \alpha, \lambda_\beta < \lambda_\alpha$, $\rightarrow$ The perturbation along $\mathbf{c}^\beta$ decays.

Case 2: $\beta \neq \alpha, \lambda_\beta = \lambda_\alpha$, $\rightarrow$ The perturbation along $\mathbf{c}^\beta$ persists.

Case 3: $\beta \neq \alpha, \lambda_\beta > \lambda_\alpha$, $\rightarrow$ The perturbation along $\mathbf{c}^\beta$ grows.

Case 4: $\beta = \alpha$, $\rightarrow$ The perturbation along $\mathbf{c}^\alpha$ always decays.

The weight vector $\mathbf{w} = \mathbf{c}^\alpha$ is stable only if perturbations in all directions decay, i.e.

$$\mathbf{w} = \mathbf{c}^\alpha \text{ stable} \iff \lambda_\beta < \lambda_\alpha \quad \forall \beta \neq \alpha\,. \tag{72}$$

Only the eigenvector $\mathbf{c}^1$ with largest eigenvalue λ_1 is a stable weight vector under Oja's rule.

Which weight vectors are stable?

$$\mathbf{w} \overset{(51)}{:=} \mathbf{c}^\alpha + \epsilon\,,$$

$$\langle \Delta s_\beta \rangle_\mu \overset{(71)}{\approx} \kappa_{\alpha\beta} s_\beta \quad \text{with} \quad \kappa_{\alpha\beta} \overset{(70)}{:=} \begin{cases} \eta(-2\lambda_\beta) & \text{if } \beta = \alpha \\ \eta(\lambda_\beta - \lambda_\alpha) & \text{if } \beta \neq \alpha \end{cases} .$$

Case 1: $\beta \neq \alpha, \lambda_\beta < \lambda_\alpha$, $\rightarrow$ The perturbation along $\mathbf{c}^\beta$ decays.

Case 2: $\beta \neq \alpha, \lambda_\beta = \lambda_\alpha$, $\rightarrow$ The perturbation along $\mathbf{c}^\beta$ persists.

Case 3: $\beta \neq \alpha, \lambda_\beta > \lambda_\alpha$, $\rightarrow$ The perturbation along $\mathbf{c}^\beta$ grows.

Case 4: $\beta = \alpha$, $\rightarrow$ The perturbation along $\mathbf{c}^\alpha$ always decays.

The weight vector $\mathbf{w} = \mathbf{c}^\alpha$ is stable only if perturbations in all directions decay, i.e.

$$\mathbf{w} = \mathbf{c}^\alpha \text{ stable} \iff \lambda_\beta < \lambda_\alpha \quad \forall \beta \neq \alpha\,. \tag{72}$$

Only the eigenvector $\mathbf{c}^1$ with largest eigenvalue λ_1 is a stable weight vector under Oja's rule.

What happens if the two largest eigenvalues are equal?

Summary

$$y \overset{(4)}{=} \mathbf{w}^T\mathbf{x} \quad \text{(linear unit)}$$

$$\Delta\mathbf{w} \overset{(20)}{=} \eta(y\mathbf{x} - y^2\mathbf{w}) \quad \text{(Oja's rule)}$$

---

1. What is the mean dynamics of the weight vectors?

$$\frac{1}{\eta}\langle\Delta\mathbf{w}\rangle_\mu \overset{(32)}{=} \mathbf{C}\mathbf{w} - \left(\mathbf{w}^T\mathbf{C}\mathbf{w}\right)\mathbf{w}$$

$$\text{with} \quad \mathbf{C} \overset{(31)}{:=} \langle\mathbf{x}\mathbf{x}^T\rangle_\mu$$

2. What are the stationary weight vectors of the dynamics?

$$\mathbf{0} \overset{!}{=} \langle\Delta\mathbf{w}\rangle_\mu \iff \mathbf{w} = \mathbf{c}^\alpha$$

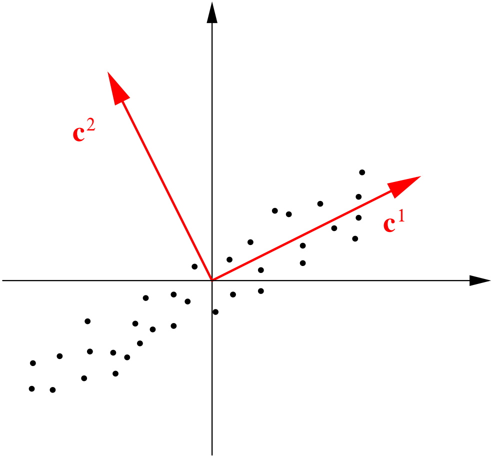
$$\text{with } \mathbf{C}\mathbf{c}^\alpha \overset{(36)}{=} \lambda_\alpha\mathbf{c}^\alpha \quad \text{(eigenvectors of } \mathbf{C}\text{)}$$

$$\wedge \quad 1 \overset{(38)}{=} \|\mathbf{c}^\alpha\|^2 \quad \text{(with norm 1)}$$

3. Which weight vectors are stable?

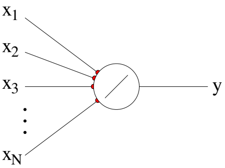
$$\mathbf{w} = \mathbf{c}^\alpha \text{ stable} \overset{(72)}{\iff} \lambda_\beta < \lambda_\alpha \quad \forall\beta \neq \alpha$$

Reminder: Principal Components



Principal components are eigenvectors of the covariance matrix and point in the direction of maximal variance within the space orthogonal to the earlier principal components.

Learning Several Principal Components

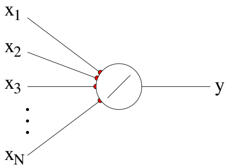


$$y \overset{(4)}{=} \mathbf{w}^T \mathbf{x} \quad \text{(linear unit)}$$

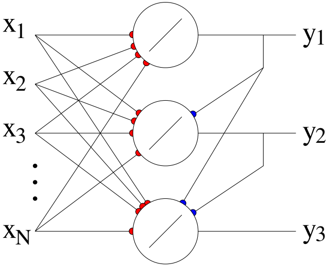
$$\Delta \mathbf{w} \overset{(20)}{=} \eta(y\mathbf{x} - y^2\mathbf{w}) \quad \text{(Oja's rule)}$$

(Rubner & Tavan, 1989, Europhys. Letters 10:693–8; reviewed in Becker & Plumbley, 1996, J. Appl. Intelligence 6:185–205)

Learning Several Principal Components



$$y \overset{(4)}{=} \mathbf{w}^T\mathbf{x} \quad \text{(linear unit)}$$
$$\Delta\mathbf{w} \overset{(20)}{=} \eta(y\mathbf{x} - y^2\mathbf{w}) \quad \text{(Oja's rule)}$$

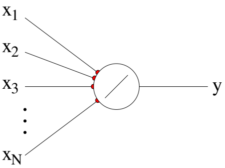


$$\begin{aligned} y_j &= \mathbf{w}_j^T\mathbf{x} + \sum_{k=1}^{j-1} v_{jk} y_k \\ \Delta\mathbf{w}_j &= \eta(y_j\mathbf{x} - y_j^2\mathbf{w}_j) \\ \Delta v_{jk} &= -\epsilon y_j y_k \end{aligned}$$

Asymmetric inhibitory lateral connections decorrelate later from earlier output units. The units learn the principal components in order of decreasing eigenvalue.

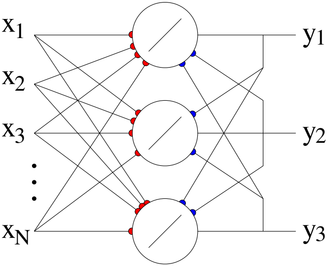
(Rubner & Tavan, 1989, Europhys. Letters 10:693–8; reviewed in Becker & Plumbley, 1996, J. Appl. Intelligence 6:185–205)

Learning a Principal Subspace



$$y \overset{(4)}{=} \mathbf{w}^T \mathbf{x} \quad \text{(linear unit)}$$

$$\Delta \mathbf{w} \overset{(20)}{=} \eta(y\mathbf{x} - y^2 \mathbf{w}) \quad \text{(Oja's rule)}$$



$$y_j = \mathbf{w}_j^T \mathbf{x} + \sum_{\substack{k=1 \\ k \neq j}}^{N} v_{jk} y_k$$

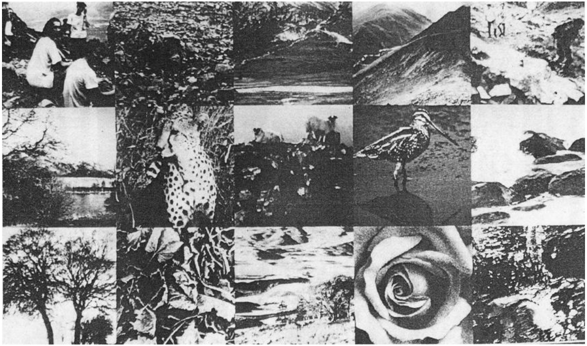
$$\Delta \mathbf{w}_j = \eta(y_j \mathbf{x} - y_j^2 \mathbf{w}_j)$$

$$\Delta v_{jk} = -\epsilon y_j y_k$$

Symmetric inhibitory lateral connections mutually decorrelate output units. The units learn the principal subspace but not particular principal components.

(Földiák, 1989, Proc. IJCNN'89 pp. 401–405; reviewed in Becker & Plumbley, 1996, J. Appl. Intelligence 6:185–205)

The Principal Components of Natural Images



15 natural images of size 256×256 pixels.

20,000 random samples of size 64×64 pixels.

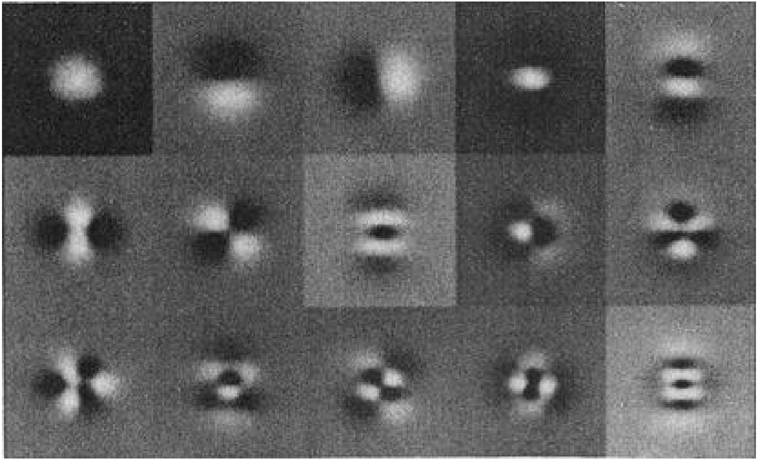
For each pixel the mean gray value over the 20,000 samples was removed.

The samples were windowed with a Gaussian with std. dev. 10 pixels.

Sanger's rule was applied to the samples.

(Hancock, Baddeley, & Smith, 1992, Network 3(1):61–70)

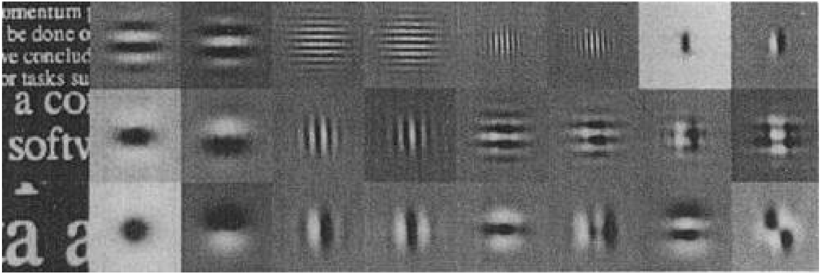
The Principal Components of Natural Images



The first principal components resemble simple-cell receptive fields in the primary visual cortex, the later ones do not.

(Hancock, Baddeley, & Smith, 1992, Network 3(1):61–70)

The Principal Components of Natural Images



The principal components look different if samples are taken from text.

(Hancock, Baddeley, & Smith, 1992, Network 3(1):61–70)