



شبکه‌های عصبی مصنوعی

جلسه بیستم:
شبکه هب (۲)
(Hebbian Network)

شبکه هب (Hebbian Network)

روش آموزش هب:

$$\Delta w_{jk} = a_2 \left[\sum_{j=1}^m w_{jk} x_k - p \right] [x_k - q] + a_3$$

شبکه هب (Hebbian Network)

روش آموزش هب:

$$\Delta w_{jk} = a_2 \left[\sum_{j=1}^m w_{jk} x_k - p \right] [x_k - q] + a_3$$

- حال، میانگین تغییرات در وزن‌ها را در اثر اعمال یک ورودی به شبکه را به دست می‌آوریم:

$$E[\Delta w_{jk}] = k_1 + \sum_{i=1}^m w_{ji} C_{ik} + \frac{k_2}{m} \sum_{i=1}^m w_{ji}$$

$$C_{ik} = E[(x_i - \bar{x})(x_k - \bar{x})]$$

شبکه هب (Hebbian Network)

روش آموزش هب:

$$\Delta w_{jk} = a_2 \left[\sum_{j=1}^m w_{jk} x_k - p \right] [x_k - q] + a_3$$

- حال، میانگین تغییرات در وزن‌ها را در اثر اعمال یک ورودی به شبکه را به دست می‌آوریم:

$$E[\Delta w_{jk}] = k_1 + \sum_{i=1}^m w_{ji} C_{ik} + \frac{k_2}{m} \sum_{i=1}^m w_{ji}$$

$$C_{ik} = E[(x_i - \bar{x})(x_k - \bar{x})]$$

- می‌توان نشان داد که تغییرات در وزن‌ها به فرم بالا، برابر است با محاسبه گرادیان نزولی

$$\Delta w_{jk} = -\eta \frac{\partial E}{\partial w_{jk}}$$

برروی تابع هزینه زیر:

$$E = -\frac{1}{2} \mathbf{w}^T \mathbf{C} \mathbf{w} + \frac{\lambda}{2} \left(\mu - \sum_{i=1}^m w_{ji} \right)^2$$

شبکه هب (Hebbian Network)

روش آموزش هب:

$$E = -\frac{1}{2} \mathbf{w}^T \mathbf{C} \mathbf{w} + \frac{\lambda}{2} \left(\mu - \sum_{i=1}^m w_{ji} \right)^2$$

شبکه هب (Hebbian Network)

روش آموزش هب:

$$E = -\frac{1}{2} \mathbf{w}^T \mathbf{C} \mathbf{w} + \frac{\lambda}{2} \left(\mu - \sum_{i=1}^m w_{ji} \right)^2$$

پیشینه کردن $\mathbf{w}^T \mathbf{C} \mathbf{w}$ (واریانس خروجی سلول ها) با قیود $w^- \leq w_{jk} \leq w^+$ و $\mu = \sum_{i=1}^m w_{ji}$

شبکه هب (Hebbian Network)

روش آموزش هب:

$$E = -\frac{1}{2} \mathbf{w}^T \mathbf{C} \mathbf{w} + \frac{\lambda}{2} \left(\mu - \sum_{i=1}^m w_{ji} \right)^2$$

پیشینه کردن $\mathbf{w}^T \mathbf{C} \mathbf{w}$ (واریانس خروجی سلول ها) با قیود $\mu = \sum_{i=1}^m w_{ji}$ و $w^- \leq w_{jk} \leq w^+$

$$\max \mathbf{w}^T \mathbf{C} \mathbf{w}$$

$$\text{s.t.} \begin{cases} \mu = \sum_{i=1}^m w_{ji} \\ w^- \leq w_{jk} \leq w^+ \end{cases}$$

شبکه هب (Hebbian Network)

روش آموزش هب:

$$E = -\frac{1}{2} \mathbf{w}^T \mathbf{C} \mathbf{w} + \frac{\lambda}{2} \left(\mu - \sum_{i=1}^m w_{ji} \right)^2$$

بیشینه کردن $\mathbf{w}^T \mathbf{C} \mathbf{w}$ (واریانس خروجی سلول ها) با قیود $\mu = \sum_{i=1}^m w_{ji}$ و $w^- \leq w_{jk} \leq w^+$

$$\begin{aligned} & \max \mathbf{w}^T \mathbf{C} \mathbf{w} \\ & \text{s.t.} \begin{cases} \mu = \sum_{i=1}^m w_{ji} \\ w^- \leq w_{jk} \leq w^+ \end{cases} \end{aligned}$$

- به عبارت دیگر، تحت آموزش هب، هر سلول سعی در **بیشینه کردن واریانس خروجی خود** با در نظر گرفتن قیود معینی را دارد.

شبکه هب (Hebbian Network)

روش آموزش هب:

واریانس خروجی سلول ها و اهمیت آن ها در استخراج ویژگی های الگوها:

شبکه هب (Hebbian Network)

روش آموزش هب:

واریانس خروجی سلول ها و اهمیت آن ها در استخراج ویژگی های الگوها:

– ماتریس کوواریانس (یا به عبارتی ماتریس همبستگی Correlation Matrix) را در نظر بگیرید:

$$\mathbf{C} = E[(\mathbf{x} - \bar{\mathbf{x}})(\mathbf{x}^T - \bar{\mathbf{x}})]$$

شبکه هب (Hebbian Network)

روش آموزش هب:

واریانس خروجی سلول ها و اهمیت آن ها در استخراج ویژگی های الگوها:

– ماتریس کوواریانس (یا به عبارتی ماتریس همبستگی Correlation Matrix) را در نظر بگیرید:

$$\mathbf{C} = E[(\mathbf{x} - \bar{\mathbf{x}})(\mathbf{x}^T - \bar{\mathbf{x}})]$$

بافرض صفر بودن میانگین

$$\mathbf{C} = E[\mathbf{x} \mathbf{x}^T]$$

شبکه هب (Hebbian Network)

روش آموزش هب:

واریانس خروجی سلول‌ها و اهمیت آن‌ها در استخراج ویژگی‌های الگوها:

– ماتریس کوواریانس (یا به عبارتی ماتریس همبستگی Correlation Matrix) را در نظر بگیرید:

$$\mathbf{C} = E[(\mathbf{x} - \bar{\mathbf{x}})(\mathbf{x}^T - \bar{\mathbf{x}})]$$

بافرض صفر بودن میانگین

$$\mathbf{C} = E[\mathbf{x} \mathbf{x}^T]$$

ماتریس \mathbf{C} متقارن و مثبت نیمه معین است

$$\mathbf{C} = \mathbf{C}^T$$

شبکه هب (Hebbian Network)

روش آموزش هب:

واریانس خروجی سلول‌ها و اهمیت آن‌ها در استخراج ویژگی‌های الگوها:

– ماتریس کوواریانس (یا به عبارتی ماتریس همبستگی Correlation Matrix) را در نظر بگیرید:

$$\mathbf{C} = E[(\mathbf{x} - \bar{\mathbf{x}})(\mathbf{x}^T - \bar{\mathbf{x}})]$$

بافرض صفر بودن میانگین

$$\mathbf{C} = E[\mathbf{x} \mathbf{x}^T]$$

ماتریس \mathbf{C} متقارن و مثبت نیمه معین است

$$\mathbf{C} = \mathbf{C}^T$$

– با استفاده از تبدیل متعامد مشابه (Orthogonal Similar Transformation) می‌توان این ماتریس را به بردارها و مقادیر ویژه به صورت زیر تجزیه کرد:

$$\mathbf{C} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T$$

شبکه هب (Hebbian Network)

روش آموزش هب:

واریانس خروجی سلول ها و اهمیت آن ها در استخراج ویژگی های الگوها:

شبکه هب (Hebbian Network)

روش آموزش هب:

واریانس خروجی سلول‌ها و اهمیت آن‌ها در استخراج ویژگی‌های الگوها:

- برطبق خاصیت مثبت نیمه‌معین بودن C ، تمام مقادیر ویژه مثبت‌اند.
- همچنین، چنانچه فرض کنیم C دارای m مقدار ویژه مجزا باشد، در این صورت می‌توان این مقادیر ویژه را به صورت زیر منظم کرد:

$$\lambda_1 > \lambda_2 > \dots > \lambda_m$$

شبکه هب (Hebbian Network)

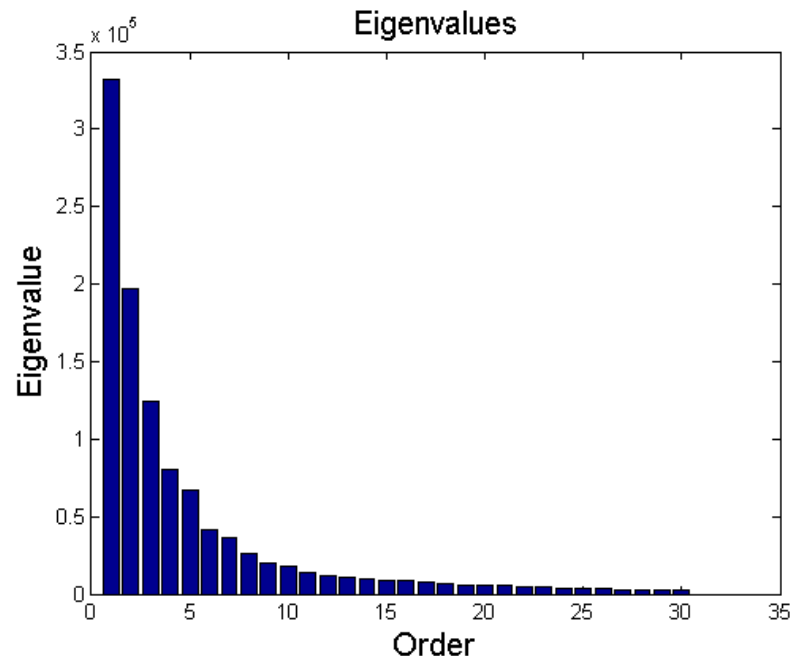
روش آموزش هب:

واریانس خروجی سلول‌ها و اهمیت آن‌ها در استخراج ویژگی‌های الگوها:

– برطبق خاصیت مثبت نیمه‌معین بودن C ، تمام مقادیر ویژه مثبت‌اند.

– همچنین، چنانچه فرض کنیم C دارای m مقدار ویژه مجزا باشد، در این صورت می‌توان این مقادیر ویژه را به صورت زیر منظم کرد:

$$\lambda_1 > \lambda_2 > \dots > \lambda_m$$



شبکه هب (Hebbian Network)

روش آموزش هب:

واریانس خروجی سلول‌ها و اهمیت آن‌ها در استخراج ویژگی‌های الگوها:

– برطبق خاصیت مثبت نیمه‌معین بودن C ، تمام مقادیر ویژه مثبت‌اند.

– همچنین، چنانچه فرض کنیم C دارای m مقدار ویژه مجزا باشد، در این صورت می‌توان این مقادیر ویژه را به صورت زیر منظم کرد:

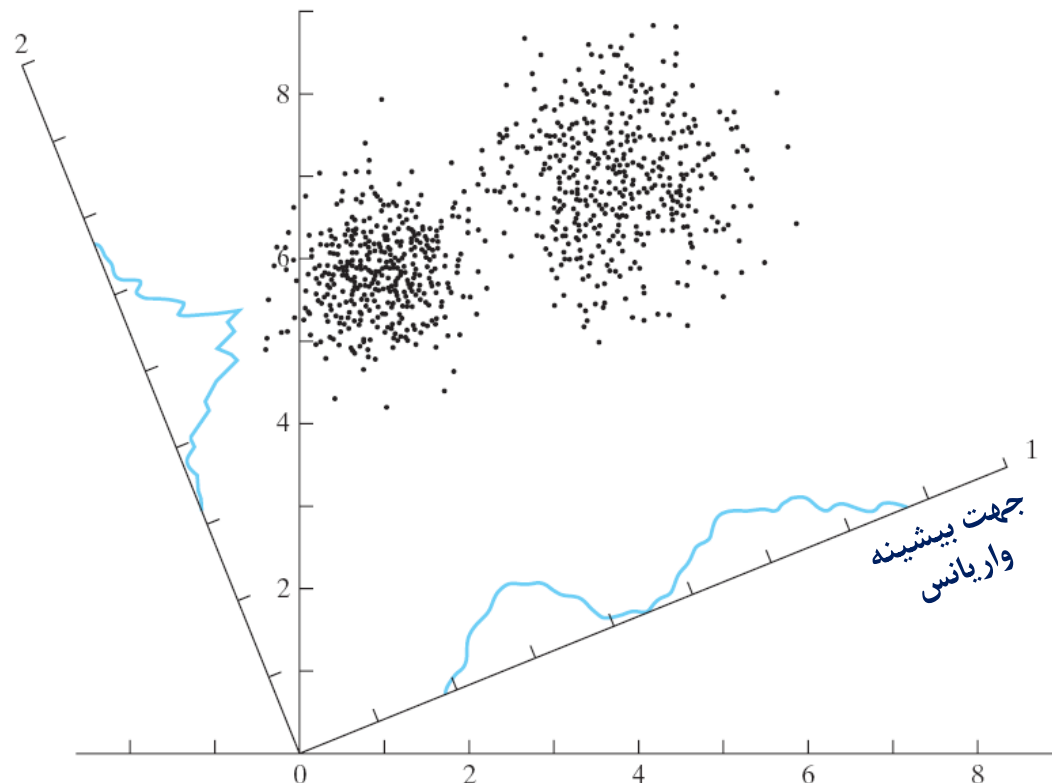
$$\lambda_1 > \lambda_2 > \dots > \lambda_m$$

– توجه کنید که این مقادیر ویژه، همان واریانس خروجی سلول‌ها هستند که از فضای داده‌ها به فضای «ویژگی داده‌ها» انتقال یافته‌اند و در این فضا، این واریانس‌ها مقدار بیشینه خود را یافته‌اند.

شبکه هب (Hebbian Network)

روش آموزش هب:

واریانس خروجی سلول‌ها و اهمیت آن‌ها در استخراج ویژگی‌های الگوها:



شبکه هب (Hebbian Network)

روش آموزش هب:

واریانس خروجی سلول‌ها و اهمیت آن‌ها در استخراج ویژگی‌های الگوها:

- بر طبق خاصیت مثبت نیمه‌معین بودن C ، تمام مقادیر ویژه مثبت‌اند.
- همچنین، چنانچه فرض کنیم C دارای m مقدار ویژه مجزا باشد، در این صورت می‌توان این مقادیر ویژه را به صورت زیر منظم کرد:

$$\lambda_1 > \lambda_2 > \dots > \lambda_m$$

- توجه کنید که این مقادیر ویژه، همان واریانس خروجی سلول‌ها هستند که از فضای داده‌ها به فضای «ویژگی داده‌ها» انتقال یافته‌اند و در این فضا، این واریانس‌ها مقدار بیشینه خود را یافته‌اند.
- مقادیر ویژه بزرگتر نماینده ویژگی‌های مهم‌تری از بردار ورودی هستند. برای درک بهتر این موضوع، تجزیه ماتریس C را به صورت زیر در نظر بگیرید:

شبکه هب (Hebbian Network)

روش آموزش هب:

واریانس خروجی سلول‌ها و اهمیت آن‌ها در استخراج ویژگی‌های الگوها:

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

شبکه هب (Hebbian Network)

روش آموزش هب:

واریانس خروجی سلول ها و اهمیت آن ها در استخراج ویژگی های الگوها:

$$\mathbf{C} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T$$

$$\mathbf{C} = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_m] \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_m^T \end{bmatrix}$$

شبکه هب (Hebbian Network)

روش آموزش هب:

واریانس خروجی سلول ها و اهمیت آن ها در استخراج ویژگی های الگوها:

$$\mathbf{C} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T$$

$$\mathbf{C} = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_m] \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_m^T \end{bmatrix}$$

$$\mathbf{C} = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \cdots + \lambda_m \mathbf{u}_m \mathbf{u}_m^T$$

شبکه هب (Hebbian Network)

روش آموزش هب:

واریانس خروجی سلول ها و اهمیت آن ها در استخراج ویژگی های الگوها:

$$\mathbf{C} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T$$

$$\mathbf{C} = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_m] \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_m^T \end{bmatrix}$$

$$\mathbf{C} = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \cdots + \lambda_m \mathbf{u}_m \mathbf{u}_m^T$$

$$\hat{\mathbf{C}} = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \cdots + \lambda_l \mathbf{u}_l \mathbf{u}_l^T \quad l < m$$

شبکه هب (Hebbian Network)

روش آموزش هب:

واریانس خروجی سلول‌ها و اهمیت آن‌ها در استخراج ویژگی‌های الگوها:

$$\mathbf{C} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T$$

$$\mathbf{C} = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_m] \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_m^T \end{bmatrix}$$

$$\mathbf{C} = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \cdots + \lambda_m \mathbf{u}_m \mathbf{u}_m^T$$

$$\hat{\mathbf{C}} = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \cdots + \lambda_l \mathbf{u}_l \mathbf{u}_l^T \quad l < m$$

- به مقادیر ویژه در پردازش تصاویر، مولفه‌های اساسی و انجام این کار را آنالیز مولفه‌های اساسی (Principal Component Analysis \equiv PCA) می‌گویند.

شبکه هب (Hebbian Network)

روش آموزش هب:

واریانس خروجی سلول‌ها و اهمیت آن‌ها در استخراج ویژگی‌های الگوها:

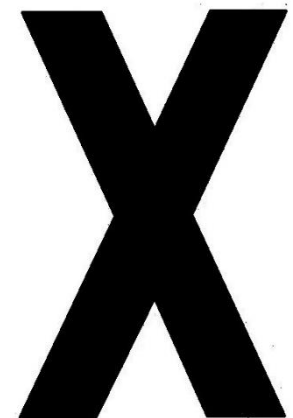
$$\hat{\mathbf{C}} = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \cdots + \lambda_l \mathbf{u}_l \mathbf{u}_l^T \quad l < m$$

شبکه هب (Hebbian Network)

روش آموزش هب:

واریانس خروجی سلول‌ها و اهمیت آن‌ها در استخراج ویژگی‌های الگوها:

$$\hat{C} = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \dots + \lambda_l \mathbf{u}_l \mathbf{u}_l^T \quad l < m$$



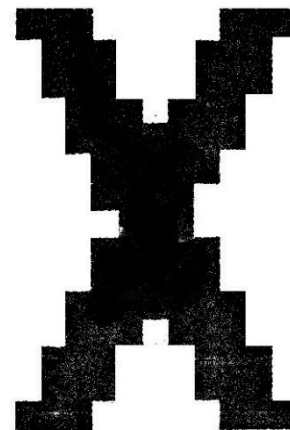
تصویر اصلی

شبکه هب (Hebbian Network)

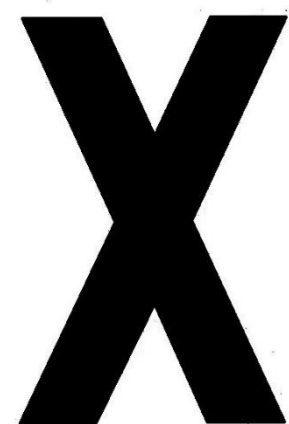
روش آموزش هب:

واریانس خروجی سلول‌ها و اهمیت آن‌ها در استخراج ویژگی‌های الگوها:

$$\hat{C} = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \dots + \lambda_l \mathbf{u}_l \mathbf{u}_l^T \quad l < m$$



تصویر گسسته شده



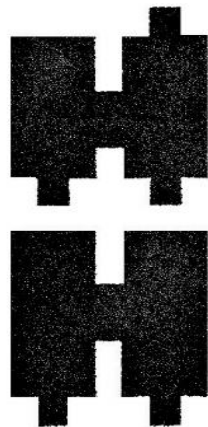
تصویر اصلی

شبکه هب (Hebbian Network)

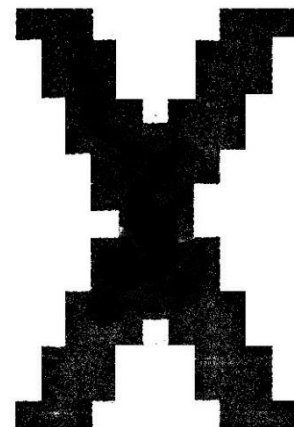
روش آموزش هب:

واریانس خروجی سلول‌ها و اهمیت آن‌ها در استخراج ویژگی‌های الگوها:

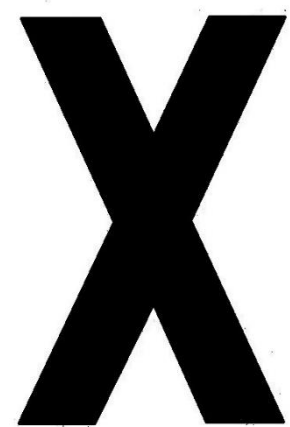
$$\hat{C} = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \dots + \lambda_l \mathbf{u}_l \mathbf{u}_l^T \quad l < m$$



بازسازی با رتبه ۱



تصویر گسسته شده



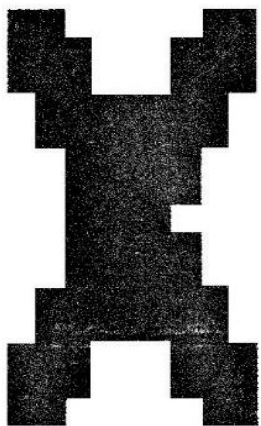
تصویر اصلی

شبکه هب (Hebbian Network)

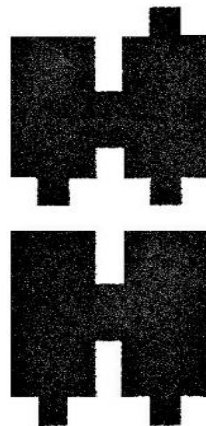
روش آموزش هب:

واریانس خروجی سلول‌ها و اهمیت آن‌ها در استخراج ویژگی‌های الگوها:

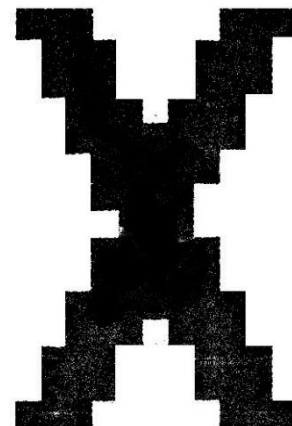
$$\hat{C} = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \dots + \lambda_l \mathbf{u}_l \mathbf{u}_l^T \quad l < m$$



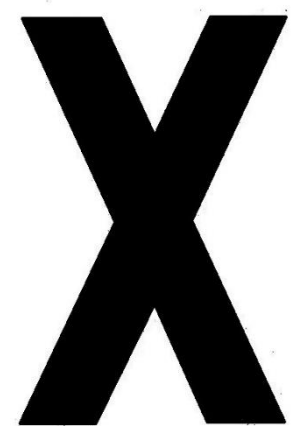
بازسازی با رتبه ۲



بازسازی با رتبه ۱



تصویر گسسته شده



تصویر اصلی

شبکه هب (Hebbian Network)

روش آموزش هب:

واریانس خروجی سلول ها و اهمیت آن ها در استخراج ویژگی های الگوها:

PCA for image compression



$p=1$



$p=2$



$p=4$



$p=8$



$p=16$



$p=32$



$p=64$



$p=100$

Original
Image



شبکه هب (Hebbian Network)

روش آموزش هب:
واریانس خروجی سلول‌ها و اهمیت آن‌ها در استخراج ویژگی‌های الگوها:

Original Image



- Divide the original 372x492 image into patches:
 - Each patch is an instance that contains 12x12 pixels on a grid
- View each as a 144-D vector

شبکه هب (Hebbian Network)

روش آموزش هب:
واریانس خروجی سلول‌ها و اهمیت آن‌ها در استخراج ویژگی‌های الگوها:

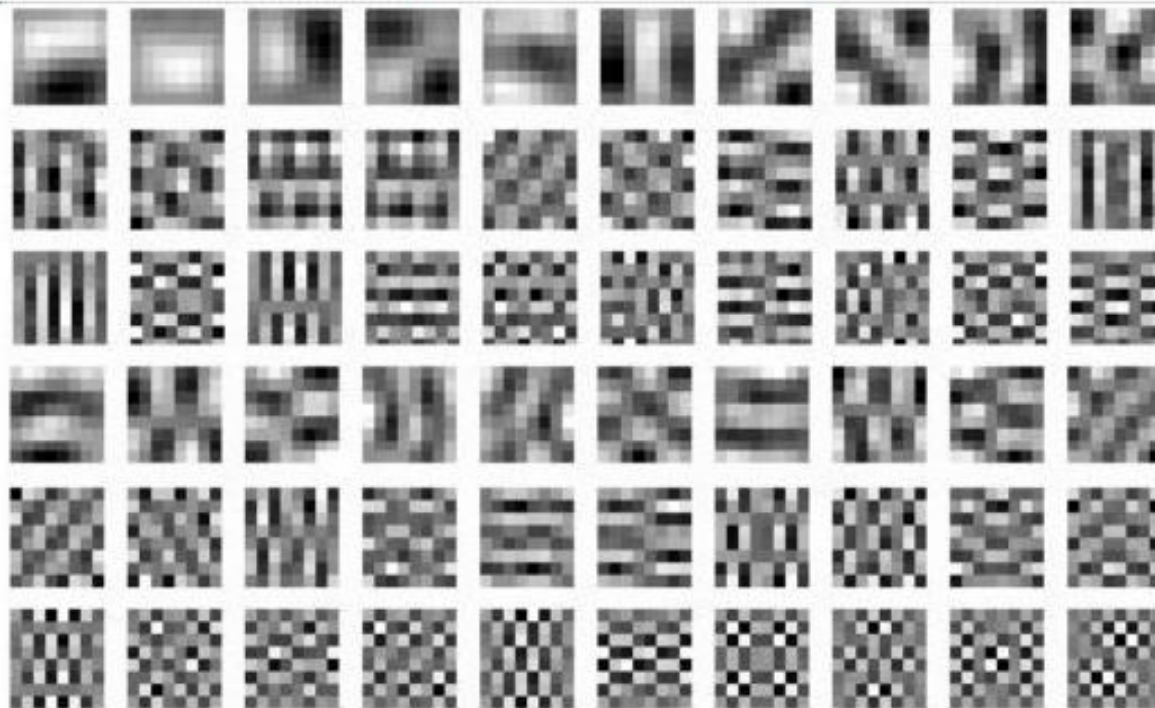
PCA compression: 144D) 60D



شبکه هب (Hebbian Network)

روش آموزش هب:
واریانس خروجی سلول ها و اهمیت آن ها در استخراج ویژگی های الگوها:

60 most important eigenvectors



شبکه هب (Hebbian Network)

روش آموزش هب:
واریانس خروجی سلول ها و اهمیت آن ها در استخراج ویژگی های الگوها:

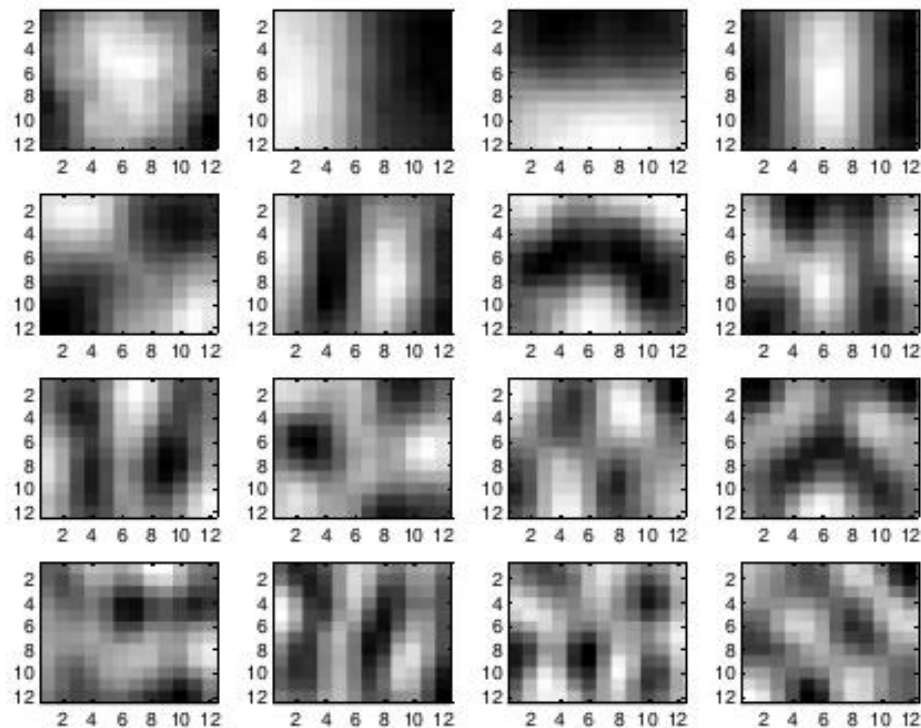
PCA compression: 144D) 16D



شبکه هب (Hebbian Network)

روش آموزش هب: واریانس خروجی سلول ها و اهمیت آن ها در استخراج ویژگی های الگوها:

16 most important eigenvectors



شبکه هب (Hebbian Network)

روش آموزش هب:
واریانس خروجی سلول ها و اهمیت آن ها در استخراج ویژگی های الگوها:

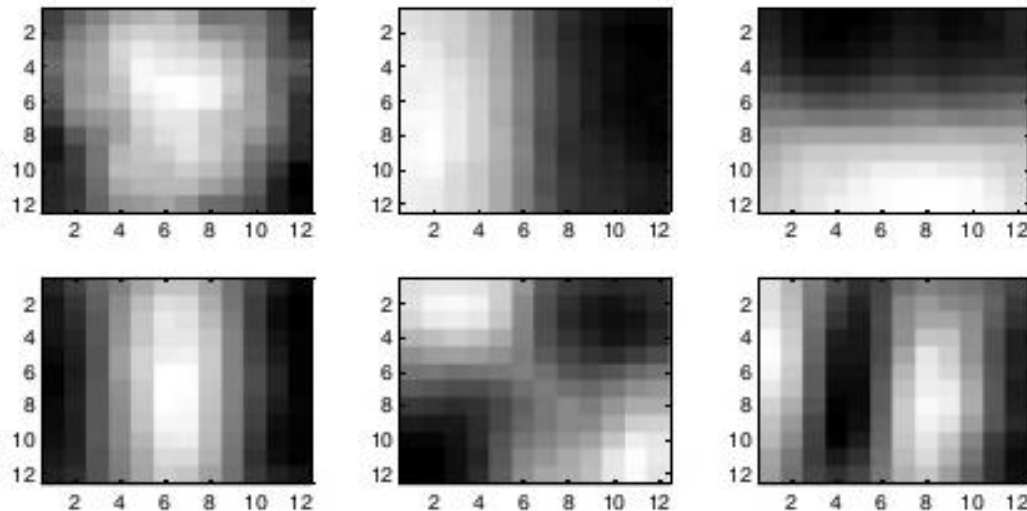
PCA compression: 144D \ 6D



شبکه هب (Hebbian Network)

روش آموزش هب:
واریانس خروجی سلول‌ها و اهمیت آن‌ها در استخراج ویژگی‌های الگوها:

6 most important eigenvectors



شبکه هب (Hebbian Network)

روش آموزش هب:
واریانس خروجی سلول‌ها و اهمیت آن‌ها در استخراج ویژگی‌های الگوها:

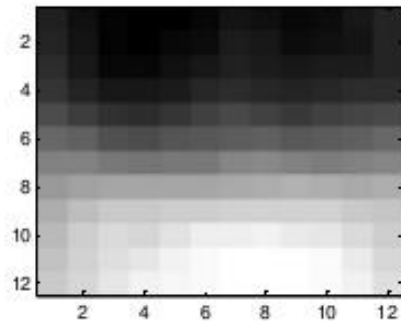
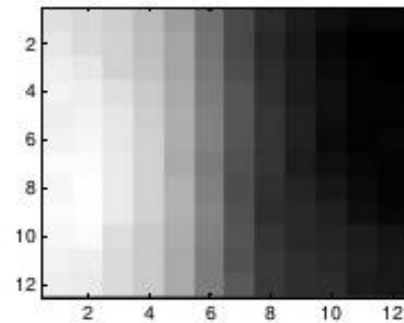
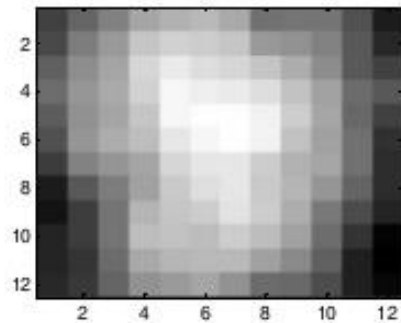
PCA compression: 144D \rightarrow 2D



شبکه هب (Hebbian Network)

روش آموزش هب: واریانس خروجی سلول ها و اهمیت آن ها در استخراج ویژگی های الگوها:

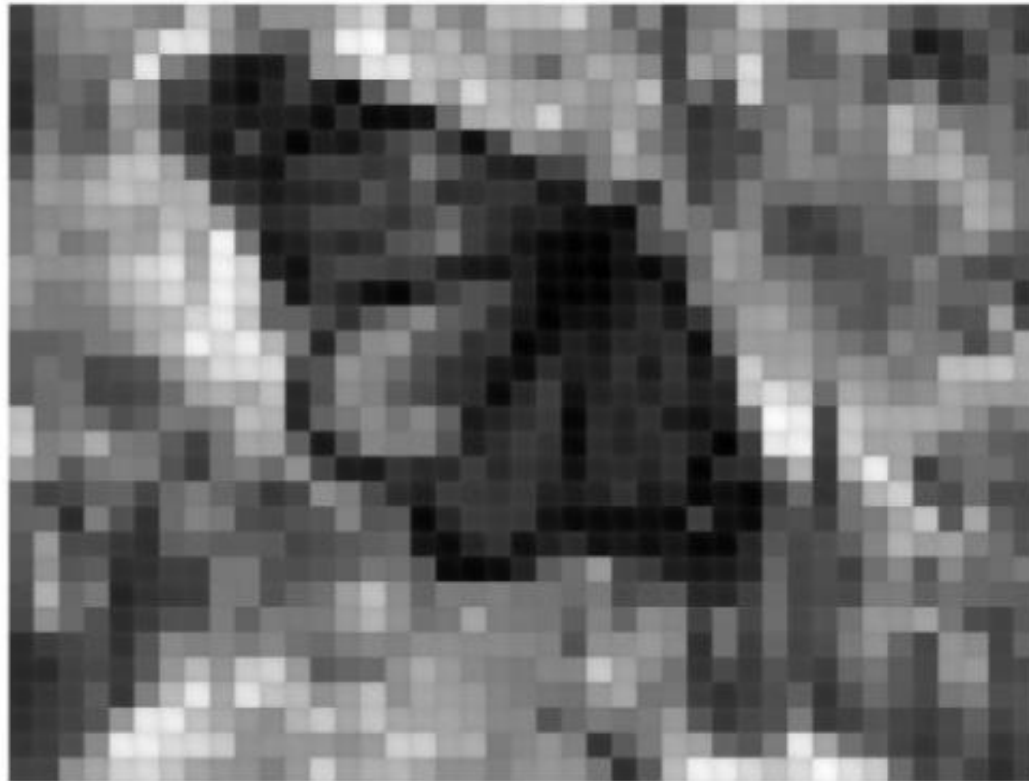
3 most important eigenvectors



شبکه هب (Hebbian Network)

روش آموزش هب: واریانس خروجی سلول ها و اهمیت آن ها در استخراج ویژگی های الگوها:

PCA compression: 144D \rightarrow 1D



شبکه هب (Hebbian Network)

روش آموزش هب: واریانس خروجی سلول ها و اهمیت آن ها در استخراج ویژگی های الگوها:

$$\hat{\mathbf{C}} = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \dots + \lambda_l \mathbf{u}_l \mathbf{u}_l^T \quad l < m$$

شبکه هب (Hebbian Network)

روش آموزش هب:
واریانس خروجی سلول‌ها و اهمیت آن‌ها در استخراج ویژگی‌های الگوها:

$$\hat{C} = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \dots + \lambda_l \mathbf{u}_l \mathbf{u}_l^T \quad l < m$$

– اکنون می‌خواهیم نشان دهیم که الگوریتم آموزش هب برای یک لایه خطی از سلول‌ها، می‌تواند همان عمل بالا را انجام دهد. یعنی استخراج بزرگترین مقدار ویژه و بردار ویژه نظیر.

شبکه هب (Hebbian Network)

روش آموزش هب: واریانس خروجی سلول ها و اهمیت آن ها در استخراج ویژگی های الگوها:

$$\hat{C} = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \dots + \lambda_l \mathbf{u}_l \mathbf{u}_l^T \quad l < m$$

– اکنون می خواهیم نشان دهیم که الگوریتم آموزش هب برای یک لایه خطی از سلول ها، می تواند همان عمل بالا را انجام دهد. یعنی استخراج بزرگترین مقدار ویژه و بردار ویژه نظیر.

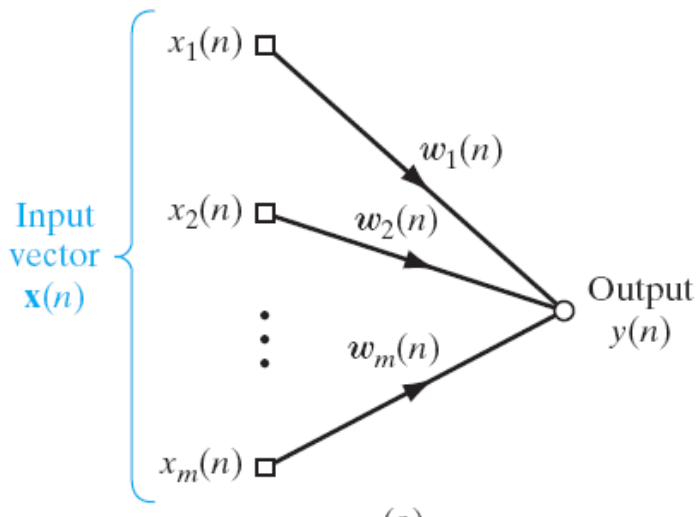
شبکه هب (Hebbian Network)

روش آموزش هب:

واریانس خروجی سلول ها و اهمیت آن ها در استخراج ویژگی های الگوها:

$$\hat{\mathbf{C}} = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \dots + \lambda_l \mathbf{u}_l \mathbf{u}_l^T \quad l < m$$

- اکنون می خواهیم نشان دهیم که الگوریتم آموزش هب برای یک لایه خطی از سلول ها، می تواند همان عمل بالا را انجام دهد. یعنی استخراج بزرگترین مقدار ویژه و بردار ویژه نظیر.



- مدل عصبی بسیار ساده زیر را در نظر بگیرید:

$$y(n) = \sum_{i=1}^m w_i(n) x_i(n)$$

شبکه هب (Hebbian Network)

روش آموزش هب:

واریانس خروجی سلول ها و اهمیت آن ها در استخراج ویژگی های الگوها:

$$\hat{\mathbf{C}} = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \dots + \lambda_l \mathbf{u}_l \mathbf{u}_l^T \quad l < m$$

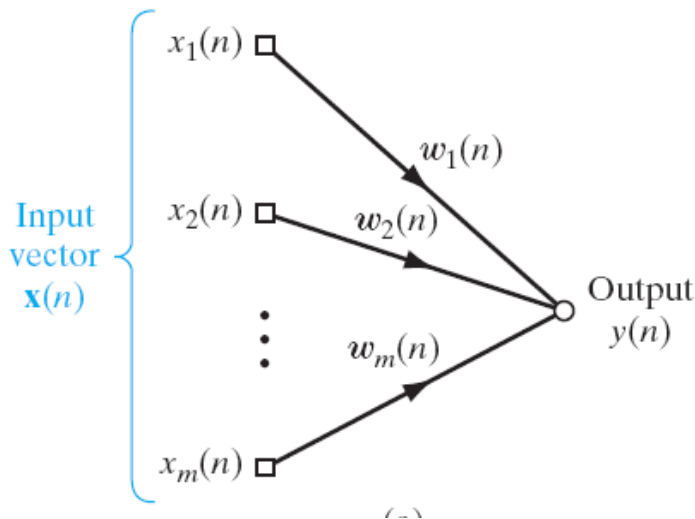
- اکنون می خواهیم نشان دهیم که الگوریتم آموزش هب برای یک لایه خطی از سلول ها، می تواند همان عمل بالا را انجام دهد. یعنی استخراج بزرگترین مقدار ویژه و بردار ویژه نظیر.

- مدل عصبی بسیار ساده زیر را در نظر بگیرید:

$$y(n) = \sum_{i=1}^m w_i(n) x_i(n)$$

- در روش آموزش هب داشتیم:

$$w_i(n+1) = w_i(n) + \eta y(n) x_i(n), \quad i = 1, \dots, m$$



شبکه هب (Hebbian Network)

روش آموزش هب:

$$y(n) = \sum_{i=1}^m w_i(n) x_i(n)$$

$$w_i(n+1) = w_i(n) + \eta y(n) x_i(n), \quad i = 1, \dots, m$$

شبکه هب (Hebbian Network)

روش آموزش هب:

$$y(n) = \sum_{i=1}^m w_i(n) x_i(n)$$

$$w_i(n+1) = w_i(n) + \eta y(n) x_i(n), \quad i = 1, \dots, m$$

– مشکل این روش آموزش این است که مقدار وزن‌ها می‌تواند بی‌رویه زیاد شود.

شبکه هب (Hebbian Network)

روش آموزش هب:

$$y(n) = \sum_{i=1}^m w_i(n) x_i(n)$$

$$w_i(n+1) = w_i(n) + \eta y(n) x_i(n), \quad i = 1, \dots, m$$

- مشکل این روش آموزش این است که مقدار وزن‌ها می‌تواند بی‌رویه زیاد شود.

- برای رفع این مشکل، از فرم خاصی از نرمال کردن وزن‌ها استفاده می‌کنیم. این کار ضمن محدود کردن مقدار وزن‌ها، باعث ایجاد رقابت بین آن‌ها نیز می‌شود:

$$w_i(n+1) = \frac{w_i(n) + \eta y(n) x_i(n)}{\left(\sum_{i=1}^m (w_i(n) + \eta y(n) x_i(n))^2 \right)^{1/2}}$$

شبکه هب (Hebbian Network)

روش آموزش هب:

$$w_i(n+1) = \frac{w_i(n) + \eta y(n)x_i(n)}{\left(\sum_{i=1}^m (w_i(n) + \eta y(n)x_i(n))^2 \right)^{1/2}}$$

شبکه هب (Hebbian Network)

روش آموزش هب:

$$w_i(n+1) = \frac{w_i(n) + \eta y(n)x_i(n)}{\left(\sum_{i=1}^m (w_i(n) + \eta y(n)x_i(n))^2 \right)^{1/2}}$$

- مخرج معادله اخیر را می توان به صورت زیر نوشت:

$$\left(\sum_{i=1}^m (w_i(n) + \eta y(n)x_i(n))^2 \right)^{1/2} = \left(\sum_{i=1}^m (w_i^2(n) + 2\eta w_i(n)y(n)x_i(n)) \right)^{1/2} + O(\eta^2)$$

- توجه کنید که در خلاصه سازی بالا از نکات زیر استفاده شده است:

- $\eta \ll 1$

شبکه هب (Hebbian Network)

روش آموزش هب:

$$w_i(n+1) = \frac{w_i(n) + \eta y(n)x_i(n)}{\left(\sum_{i=1}^m (w_i(n) + \eta y(n)x_i(n))^2 \right)^{1/2}}$$

- مخرج معادله اخیر را می توان به صورت زیر نوشت:

$$\begin{aligned} \left(\sum_{i=1}^m (w_i(n) + \eta y(n)x_i(n))^2 \right)^{1/2} &= \left(\sum_{i=1}^m (w_i^2(n) + 2\eta w_i(n)y(n)x_i(n)) \right)^{1/2} + O(\eta^2) \\ &= \left(\sum_{i=1}^m w_i^2(n) + 2\eta y(n) \sum_{i=1}^m w_i(n)x_i(n) \right)^{1/2} + O(\eta^2) \end{aligned}$$

- توجه کنید که در خلاصه سازی بالا از نکات زیر استفاده شده است:

- $\eta \ll 1$

شبکه هب (Hebbian Network)

روش آموزش هب:

$$w_i(n+1) = \frac{w_i(n) + \eta y(n)x_i(n)}{\left(\sum_{i=1}^m (w_i(n) + \eta y(n)x_i(n))^2 \right)^{1/2}}$$

- مخرج معادله اخیر را می توان به صورت زیر نوشت:

$$\begin{aligned} \left(\sum_{i=1}^m (w_i(n) + \eta y(n)x_i(n))^2 \right)^{1/2} &= \left(\sum_{i=1}^m (w_i^2(n) + 2\eta w_i(n)y(n)x_i(n)) \right)^{1/2} + O(\eta^2) \\ &= \left(\sum_{i=1}^m w_i^2(n) + 2\eta y(n) \sum_{i=1}^m w_i(n)x_i(n) \right)^{1/2} + O(\eta^2) \\ &= \left(1 + 2\eta y^2(n) \right)^{1/2} + O(\eta^2) \end{aligned}$$

- توجه کنید که در خلاصه سازی بالا از نکات زیر استفاده شده است:

- $\eta \ll 1$
- $\sum_{i=1}^m w_i^2(n) = \|\mathbf{w}(n)\|^2 = 1$
- $y(n) = \sum_{i=1}^m w_i(n)x_i(n)$

شبکه هب (Hebbian Network)

روش آموزش هب:

$$w_i(n+1) = \frac{w_i(n) + \eta y(n)x_i(n)}{\left(\sum_{i=1}^m (w_i(n) + \eta y(n)x_i(n))^2 \right)^{1/2}}$$

- مخرج معادله اخیر را می توان به صورت زیر نوشت:

$$\begin{aligned} \left(\sum_{i=1}^m (w_i(n) + \eta y(n)x_i(n))^2 \right)^{1/2} &= \left(\sum_{i=1}^m (w_i^2(n) + 2\eta w_i(n)y(n)x_i(n)) \right)^{1/2} + O(\eta^2) \\ &= \left(\sum_{i=1}^m w_i^2(n) + 2\eta y(n) \sum_{i=1}^m w_i(n)x_i(n) \right)^{1/2} + O(\eta^2) \\ &= (1 + 2\eta y^2(n))^{1/2} + O(\eta^2) \\ &= 1 + \eta y^2(n) + O(\eta^2) \end{aligned}$$

- توجه کنید که در خلاصه سازی بالا از نکات زیر استفاده شده است:

- $\eta \ll 1$
- $\sum_{i=1}^m w_i^2(n) = \|\mathbf{w}(n)\|^2 = 1$
- $y(n) = \sum_{i=1}^m w_i(n)x_i(n)$
- $(1+x)^{1/2} = 1 + x/2 - x^2/8 + x^3/16 - \dots$

شبکه هب (Hebbian Network)

روش آموزش هب:

$$w_i(n+1) = \frac{w_i(n) + \eta y(n)x_i(n)}{1 + \eta y^2(n) + O(\eta^2)}$$

– بنابراین

شبکه هب (Hebbian Network)

روش آموزش هب:

– بنابراین

$$\begin{aligned}w_i(n+1) &= \frac{w_i(n) + \eta y(n)x_i(n)}{1 + \eta y^2(n) + O(\eta^2)} \\&= (w_i(n) + \eta y(n)x_i(n)) (1 + \eta y^2(n) + O(\eta^2))^{-1}\end{aligned}$$

شبکه هب (Hebbian Network)

روش آموزش هب:

– بنابراین

$$\begin{aligned}w_i(n+1) &= \frac{w_i(n) + \eta y(n)x_i(n)}{1 + \eta y^2(n) + O(\eta^2)} \\&= (w_i(n) + \eta y(n)x_i(n)) (1 + \eta y^2(n) + O(\eta^2))^{-1} \\&= (w_i(n) + \eta y(n)x_i(n)) (1 - \eta y^2(n) + O(\eta^2))\end{aligned}$$

شبکه هب (Hebbian Network)

روش آموزش هب:

- بنابراین

$$\begin{aligned}w_i(n+1) &= \frac{w_i(n) + \eta y(n)x_i(n)}{1 + \eta y^2(n) + O(\eta^2)} \\&= (w_i(n) + \eta y(n)x_i(n)) (1 + \eta y^2(n) + O(\eta^2))^{-1} \\&= (w_i(n) + \eta y(n)x_i(n)) (1 - \eta y^2(n) + O(\eta^2))\end{aligned}$$

- برای این خلاصه سازی:

$$(1+x)^{-1} = 1 - x + x^2 - x^3 + \dots$$

شبکه هب (Hebbian Network)

روش آموزش هب:

- بنابراین

$$\begin{aligned}w_i(n+1) &= \frac{w_i(n) + \eta y(n)x_i(n)}{1 + \eta y^2(n) + O(\eta^2)} \\&= (w_i(n) + \eta y(n)x_i(n)) (1 + \eta y^2(n) + O(\eta^2))^{-1} \\&= (w_i(n) + \eta y(n)x_i(n)) (1 - \eta y^2(n) + O(\eta^2)) \\&= w_i(n) + \eta y(n)x_i(n) - \eta y^2(n)w_i(n) + O(\eta^2)\end{aligned}$$

- برای این خلاصه سازی:

$$(1+x)^{-1} = 1 - x + x^2 - x^3 + \dots$$

شبکه هب (Hebbian Network)

روش آموزش هب:

- بنابراین

$$\begin{aligned}w_i(n+1) &= \frac{w_i(n) + \eta y(n)x_i(n)}{1 + \eta y^2(n) + O(\eta^2)} \\&= (w_i(n) + \eta y(n)x_i(n)) (1 + \eta y^2(n) + O(\eta^2))^{-1} \\&= (w_i(n) + \eta y(n)x_i(n)) (1 - \eta y^2(n) + O(\eta^2)) \\&= w_i(n) + \eta y(n)x_i(n) - \eta y^2(n)w_i(n) + O(\eta^2)\end{aligned}$$

- برای این خلاصه سازی:

$$(1+x)^{-1} = 1 - x + x^2 - x^3 + \dots$$

- با صرف نظر کردن از جملات درجه بالا، نتیجه می شود:

$$w_i(n+1) = w_i(n) + \eta y(n)(x_i(n) - y(n)w_i(n))$$

شبکه هب (Hebbian Network)

روش آموزش هب:

- بنابراین

$$\begin{aligned}w_i(n+1) &= \frac{w_i(n) + \eta y(n)x_i(n)}{1 + \eta y^2(n) + O(\eta^2)} \\&= (w_i(n) + \eta y(n)x_i(n)) (1 + \eta y^2(n) + O(\eta^2))^{-1} \\&= (w_i(n) + \eta y(n)x_i(n)) (1 - \eta y^2(n) + O(\eta^2)) \\&= w_i(n) + \eta y(n)x_i(n) - \eta y^2(n)w_i(n) + O(\eta^2)\end{aligned}$$

- برای این خلاصه سازی:

$$(1+x)^{-1} = 1 - x + x^2 - x^3 + \dots$$

- با صرف نظر کردن از جملات درجه بالا، نتیجه می شود:

$$w_i(n+1) = w_i(n) + \eta y(n)(x_i(n) - y(n)w_i(n))$$

- مقایسه کنید با:

$$w_i(n+1) = w_i(n) + \eta y(n)x_i(n)$$

شبکه هب (Hebbian Network)

روش آموزش هب:

$$w_i(n+1) = w_i(n) + \eta y(n)(x_i(n) - y(n)w_i(n))$$

شبکه هب (Hebbian Network)

روش آموزش هب:

$$w_i(n+1) = w_i(n) + \eta y(n) (x_i(n) - y(n)w_i(n))$$

– با تعریف

$$x'_i(n) := x_i(n) - y(n)w_i(n)$$

شبکه هب (Hebbian Network)

روش آموزش هب:

$$w_i(n+1) = w_i(n) + \eta y(n) (x_i(n) - y(n)w_i(n))$$

– با تعریف

$$x'_i(n) := x_i(n) - y(n)w_i(n)$$

آموزش هب به صورت زیر درمی آید:

$$w_i(n+1) = w_i(n) + \eta y(n) x'_i(n)$$

شبکه هب (Hebbian Network)

روش آموزش هب:

$$w_i(n+1) = w_i(n) + \eta y(n)(x_i(n) - y(n)w_i(n))$$

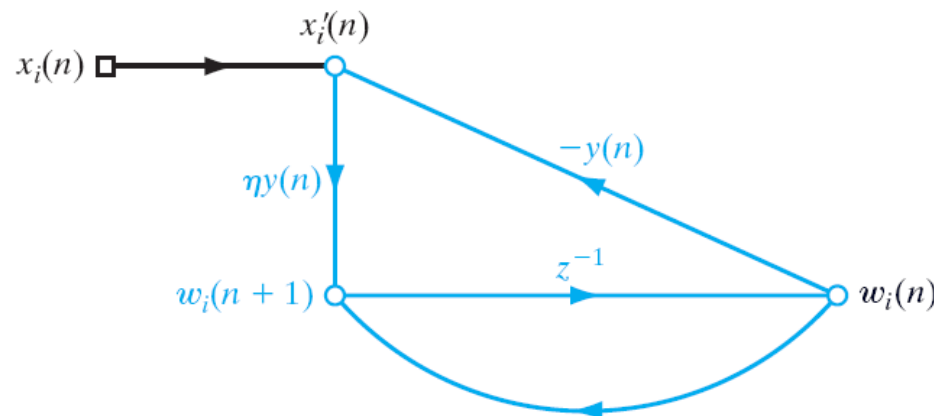
- با تعریف

$$x'_i(n) := x_i(n) - y(n)w_i(n)$$

آموزش هب به صورت زیر درمی آید:

$$w_i(n+1) = w_i(n) + \eta y(n)x'_i(n)$$

که نمودار گذر سیگنال آن به این شکل درمی آید:



شبکه هب (Hebbian Network)

شبکه هب (Hebbian Network)

همگرایی در آموزش هب:

شبکه هب (Hebbian Network)

همگرایی در آموزش هب:

$$w_i(n+1) = w_i(n) + \eta y(n)(x_i(n) - y(n)w_i(n)), \quad i = 1, \dots, m$$

شبکه هب (Hebbian Network)

همگرایی در آموزش هب:

$$w_i(n+1) = w_i(n) + \eta y(n)(x_i(n) - y(n)w_i(n)), \quad i = 1, \dots, m$$

– فرم برداری این معادله

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \eta y(n)[\mathbf{x}(n) - y(n)\mathbf{w}(n)]$$

شبکه هب (Hebbian Network)

همگرایی در آموزش هب:

$$w_i(n+1) = w_i(n) + \eta y(n)(x_i(n) - y(n)w_i(n)), \quad i = 1, \dots, m$$

– فرم برداری این معادله

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \eta y(n)[\mathbf{x}(n) - y(n)\mathbf{w}(n)]$$

– با استفاده از

$$y(n) = \mathbf{w}^T(n)\mathbf{x}(n) = \mathbf{x}^T(n)\mathbf{w}(n)$$

شبکه هب (Hebbian Network)

همگرایی در آموزش هب:

$$w_i(n+1) = w_i(n) + \eta y(n) (x_i(n) - y(n)w_i(n)), \quad i = 1, \dots, m$$

– فرم برداری این معادله

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \eta y(n) [\mathbf{x}(n) - y(n)\mathbf{w}(n)]$$

– با استفاده از

$$y(n) = \mathbf{w}^T(n)\mathbf{x}(n) = \mathbf{x}^T(n)\mathbf{w}(n)$$

نتیجه می شود:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \eta [\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{w}(n) - \mathbf{w}^T(n)\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{w}(n)\mathbf{w}(n)]$$

شبکه هب (Hebbian Network)

همگرایی در آموزش هب:

– می توان نشان داد که این الگوریتم، برای تکرارهای زیاد ($n \rightarrow \infty$) تحت سه شرط زیر همگرا می شود:

شبکه هب (Hebbian Network)

همگرایی در آموزش هب:

- می توان نشان داد که این الگوریتم، برای تکرارهای زیاد ($n \rightarrow \infty$) تحت سه شرط زیر همگرا می شود:

۱- ضریب آموزش به اندازه کافی کوچک باشد به طوری که بتوان گفت

$$E[\mathbf{w}(n+1) | \mathbf{w}(n)] = \mathbf{w}(n) + \Delta \mathbf{w}(n)$$

شبکه هب (Hebbian Network)

همگرایی در آموزش هب:

- می توان نشان داد که این الگوریتم، برای تکرارهای زیاد ($n \rightarrow \infty$) تحت سه شرط زیر همگرا می شود:

۱- ضریب آموزش به اندازه کافی کوچک باشد به طوری که بتوان گفت

$$E[\mathbf{w}(n+1) | \mathbf{w}(n)] = \mathbf{w}(n) + \Delta \mathbf{w}(n)$$

۲- ماتریس C دارای مقادیر ویژه مجزا باشد.

شبکه هب (Hebbian Network)

همگرایی در آموزش هب:

- می توان نشان داد که این الگوریتم، برای تکرارهای زیاد ($n \rightarrow \infty$) تحت سه شرط زیر همگرا می شود:

۱- ضریب آموزش به اندازه کافی کوچک باشد به طوری که بتوان گفت

$$E[\mathbf{w}(n+1) | \mathbf{w}(n)] = \mathbf{w}(n) + \Delta \mathbf{w}(n)$$

۲- ماتریس C دارای مقادیر ویژه مجزا باشد.

۳- بردار ورودی $\mathbf{x}(n)$ و بردار وزن $\mathbf{w}(n)$ از نظر آماری از یکدیگر مستقل باشند.

شبکه هب (Hebbian Network)

همگرایی در آموزش هب:

- می توان نشان داد که این الگوریتم، برای تکرارهای زیاد ($n \rightarrow \infty$) تحت سه شرط زیر همگرا می شود:

۱- ضریب آموزش به اندازه کافی کوچک باشد به طوری که بتوان گفت

$$E[\mathbf{w}(n+1) | \mathbf{w}(n)] = \mathbf{w}(n) + \Delta \mathbf{w}(n)$$

۲- ماتریس C دارای مقادیر ویژه مجزا باشد.

۳- بردار ورودی $\mathbf{x}(n)$ و بردار وزن $\mathbf{w}(n)$ از نظر آماری از یکدیگر مستقل باشند.

شبکه هب (Hebbian Network)

همگرایی در آموزش هب:

- می توان نشان داد که این الگوریتم، برای تکرارهای زیاد ($n \rightarrow \infty$) تحت سه شرط زیر همگرا می شود:

۱- ضریب آموزش به اندازه کافی کوچک باشد به طوری که بتوان گفت

$$E[\mathbf{w}(n+1) | \mathbf{w}(n)] = \mathbf{w}(n) + \Delta \mathbf{w}(n)$$

۲- ماتریس C دارای مقادیر ویژه مجزا باشد.

۳- بردار ورودی $\mathbf{x}(n)$ و بردار وزن $\mathbf{w}(n)$ از نظر آماری از یکدیگر مستقل باشند.

- چنانچه قضیه بالا را بپذیریم، در این صورت:

$$\mathbf{w}(n) \xrightarrow[n \rightarrow \infty]{} \mathbf{q}_o$$

شبکه هب (Hebbian Network)

همگرایی در آموزش هب:

- می توان نشان داد که این الگوریتم، برای تکرارهای زیاد ($n \rightarrow \infty$) تحت سه شرط زیر همگرا می شود:

۱- ضریب آموزش به اندازه کافی کوچک باشد به طوری که بتوان گفت

$$E[\mathbf{w}(n+1) | \mathbf{w}(n)] = \mathbf{w}(n) + \Delta \mathbf{w}(n)$$

۲- ماتریس C دارای مقادیر ویژه مجزا باشد.

۳- بردار ورودی $\mathbf{x}(n)$ و بردار وزن $\mathbf{w}(n)$ از نظر آماری از یکدیگر مستقل باشند.

- چنانچه قضیه بالا را بپذیریم، در این صورت:

$$\mathbf{w}(n) \xrightarrow[n \rightarrow \infty]{} \mathbf{q}_o$$

- حال می خواهیم نشان دهیم که این همگرایی به بردار ویژه ماتریس C است.