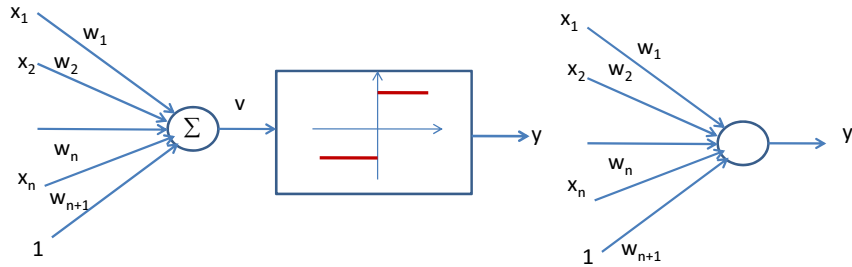


Genlikte Ayırık Algılayıcı-GAA (Perceptron)



$$v = [w_1 \ w_2 \ \dots \ w_n \ w_{n+1}] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ 1 \end{bmatrix}$$

$$v = w_1 x_1 + w_2 x_2 + \dots + w_n x_n + w_{n+1} \cdot 1$$

$$y = \varphi(v) = \begin{cases} 1 & v \geq 0 \\ -1 & v < 0 \end{cases}$$

Öğrenme Kuralı:

$$\begin{aligned} x \in S_1 & \quad w^T x > 0 & w(k+1) &= w(k) \\ x \in S_2 & \quad w^T x \leq 0 & w(k+1) &= w(k) \\ x \in S_1 & \quad w^T x \leq 0 & w(k+1) &= w(k) + cx \\ x \in S_2 & \quad w^T x > 0 & w(k+1) &= w(k) - cx \end{aligned}$$

öğrenme hızı <1 olan pozitif bir sayı

$$\rightarrow \Delta w_i = c \frac{1}{2} [y_d - y] x_i$$

$$\Delta w = c \frac{1}{2} [y_d - y] x$$

Genlikte Ayırık Algılayıcı Yakınsama Teoremi

Eğitim kümesi doğrusal ayrıştırılabilir ise genlikte ayırık algılayıcı için verilen öğrenme kuralı sonlu adımda bir çözüm verir.

Tanıt: $c=1$

X eğitim kümesindeki tüm girişler olsun

$$x^{(1)}, x^{(2)}, \dots, x^{(k)}, \dots$$

Bu eğitim kümesine karşı düşen ağırlık vektörleri

$$w^{(1)}, w^{(2)}, \dots, w^{(k)}, \dots$$

$w^{(k+1)} = w^{(k)}$ olan adımlarda eğitim kümesindeki örüntüler ve ağırlık vektörleri yukarıdaki dizilerden çıkarılsın

Ağırlıkların güncellenmesi $w^{(j)T} x^{(j)} \leq 0$ ve $x^{(j)} \in S_1$ gereken her adımda ne $w^{(j)T} x^{(j)} > 0$ ve $x^{(j)} \in S_2$ oluyor?

Güncelleme sırasında neler olacak

$$w^{(2)} = w^{(1)} + x^{(1)}$$

$$w^{(3)} = w^{(2)} + x^{(2)} = w^{(1)} \oplus x^{(1)} \oplus x^{(2)}$$

$$w^{(4)} = w^{(3)} + x^{(3)} = w^{(1)} \oplus x^{(1)} \oplus x^{(2)} \oplus x^{(3)}$$

$$\vdots$$

$$w^{(k+1)} = w^{(k)} + x^{(k)} = w^{(1)} \oplus x^{(1)} \oplus \dots \oplus x^{(k)}$$

öğrenme kuralı

$$\begin{aligned} x \in S_1 & \quad w^T x \leq 0 & w(k+1) &= w(k) \oplus x \\ x \in S_2 & \quad w^T x > 0 & w(k+1) &= w(k) \ominus x \end{aligned}$$

Burada tuhaf olan birşey var.

Geriye hangileri kaldı?

Eğitim kümesi doğrusal ayrıştırılabilir

$$\rightarrow \exists \Omega \text{ çözüm kümesi } \ni w \in \Omega, w^T x > 0$$

$$\rightarrow \text{Bir } w^* \in \Omega \text{ seçilsin } \alpha \hat{=} \min_{x \in X} w^{*T} x$$

Burada da tuhaf olan birşey var.

$$\beta \hat{=} w^{*T} w^{(1)}$$

$$w^{*T} w^{(k+1)} = \underbrace{w^{*T} w^{(1)}}_{\beta} + \underbrace{w^{*T} x^{(1)}}_{\geq \alpha} + \dots + \underbrace{w^{*T} x^{(k)}}_{\geq \alpha}$$

$$w^{*T} w^{(k+1)} \geq \beta + k\alpha$$

$$\|w^{*T}\|^2 \|w^{(k+1)}\|^2 \geq \|w^{*T} w^{(k+1)}\|^2 \geq (\beta + k\alpha)^2$$

$$\|w^{(k+1)}\|^2 \geq \frac{(k\alpha + \beta)^2}{\|w^*\|^2}$$

Alt sınır

$$w^{(i+1)} = w^{(i)} + x^{(i)}$$

$$\|w^{(i+1)}\|^2 = w^{(i+1)T} w^{(i+1)} = w^{(i)T} w^{(i)} + 2w^{(i)T} x^{(i)} + x^{(i)T} x^{(i)}$$

$$= \|w^{(i)}\|^2 + 2w^{(i)T} x^{(i)} + \|x^{(i)}\|^2$$

Neden?

$$w^{(i)T} x^{(i)} \leq 0 \rightarrow \|w^{(i+1)}\|^2 - \|w^{(i)}\|^2 \leq \|x^{(i)}\|^2$$

Nasıl geldik?

$$A \hat{=} \max_{x \in X} \|x\|^2 \rightarrow \|w^{(k+1)}\|^2 \leq \sum_{i=1}^k \|x^{(i)}\|^2 + \|w^{(1)}\|^2$$

$$\leq kA + \|w^{(1)}\|^2$$

Üst sınır

$$\|w^{(k+1)}\|^2 \geq \frac{(k\alpha + \beta)^2}{\|w^*\|^2} \quad \|w^{(k+1)}\|^2 \leq kA + \|w^{(1)}\|^2$$

Alt sınır

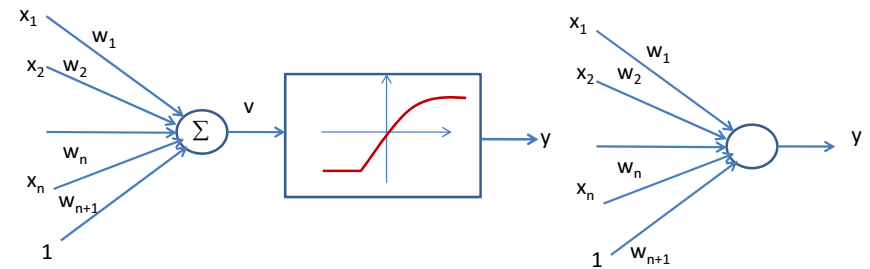
Üst sınır

Büyük k'lar için bu sınırlara dikkat edin!!!

$$\frac{(k_m \alpha + \beta)^2}{\|w^*\|^2} = k_m A + \|w^{(1)}\|^2$$

en büyük adım sayısı yukarıdaki eşitliğin belirlediği k_m 'den büyük olamaz ■

Genlikte Sürekli Algılayıcı (ADALİNE)



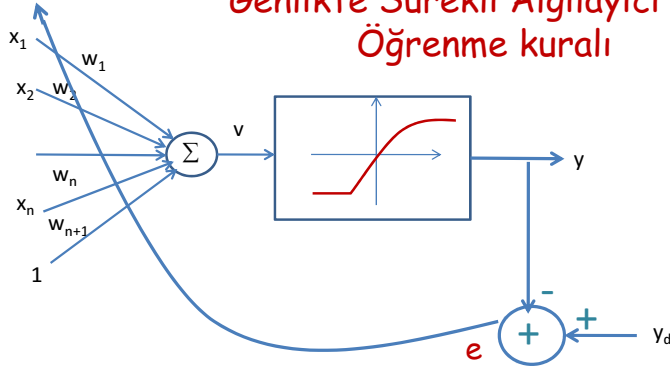
$$v = [w_1 \ w_2 \ \dots \ w_n \ w_{n+1}] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$v = w_1 x_1 + w_2 x_2 + \dots + w_n x_n + w_{n+1} 1$$

$$y = \phi(v) = \tanh av$$

$$y = \phi(v) = \frac{1}{1 + e^{-av}}$$

Genlikte Sürekli Algılayıcı için Öğrenme kuralı



Amaç: hatayı azaltacak ağırlıkları belirlemek

Hataya ilişkin bir fonksiyon oluşturularak işe başlanacak

$$e = y_d - y$$

Neden bir fonksiyon?

$$E = \frac{1}{2} e^T e = (y_d - y)^T (y_d - y) = (y_d - \phi(v))^T (y_d - \phi(v))$$

Nasıl azaltabiliriz?

Amaç: $\min_{w \in R^{n+1}} E$ Neden vazgeçtik?

$$\begin{aligned} E &= \frac{1}{2} e^T e = (y_d - y)^T (y_d - y) = (y_d - \phi(v))^T (y_d - \phi(v)) \\ &= (y_d - \phi(w_1 x_1 + w_2 x_2 + \dots + w_n x_n + w_{n+1} 1))^T \\ &\quad (y_d - \phi(w_1 x_1 + w_2 x_2 + \dots + w_n x_n + w_{n+1} 1)) \end{aligned}$$

Böylece E'nin w'ya bağımlılığını açıkça yazdık, acaba $\min_{w \in R^{n+1}} E$ sağlayan w'ları nasıl buluruz?

$E(w^{(k+1)}) \leq E(w^{(k)})$ sağlayacak $w^{(k+1)}$ 'i $w^{(k)}$ 'dan nasıl elde ederiz?

$$w^{(k)} + ? = w^{(k+1)} \rightarrow E(w^{(k+1)}) \leq E(w^{(k)})$$

$$w^{(k)} - c \nabla E = w^{(k+1)}$$

Nasıl bulunur?

Aktivasyon fonksiyonunu lineer alırsak ...

$$\Delta w = -(y_d - y) \phi'(v) x \rightarrow \Delta w = -(y_d - y) x$$

Bir de girişleri normalize edersek ...

$$\Delta w = -(y_d - y) \phi'(v) x \rightarrow \Delta w = -(y_d - y) \frac{x}{\|x\|}$$

ADALİNE'ı nerede kullanabiliriz

• sınıflandırma probleminde

$$\text{Eğitim aşamasından sonra } y = \begin{cases} 1 & \sum_{i=1}^n w_{ij} x_i > 0 \\ -1 & \sum_{i=1}^n w_{ij} x_i \leq 0 \end{cases}$$

• yaklaşık eğri uydurma

• Lineer regresyon

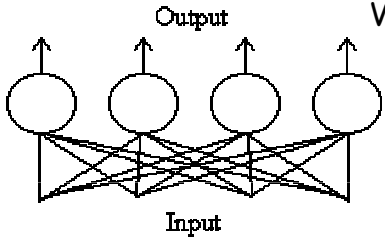
$$\nabla E \triangleq \begin{bmatrix} \frac{\partial E}{\partial w_1} \\ \frac{\partial E}{\partial w_2} \\ \vdots \\ \frac{\partial E}{\partial w_{n+1}} \end{bmatrix} = \frac{\partial E}{\partial e} \frac{\partial e}{\partial y} \frac{\partial y}{\partial v} \frac{\partial v}{\partial w} = -(y_d - y) \phi'(v) \begin{bmatrix} \frac{\partial v}{\partial w_1} \\ \frac{\partial v}{\partial w_2} \\ \vdots \\ \frac{\partial v}{\partial w_{n+1}} \end{bmatrix}$$

Öğrenme Kuralı:

$$\Delta w = -(y_d - y) \phi'(v) x = -(y_d - y) \phi'(v) \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ 1 \end{bmatrix}$$

Hep iki sınıfa ayırdık daha fazla sınıfa nasıl ayıracağız?

Unutulmaması gereken kısıt ne?



<http://richardbowles.tripod.com/neural/slpmlp/>

Verilenler: $\left\{ \left(x^l, y_d^l \right) \right\}_{l=1}^P$ Eğitim Kümesi

Amaç: m sınıfa ayırmak

$$S_1 : \sum_{i=1}^n w_i x_i + w_{n+1} < T_1$$

$$S_2 : T_1 \leq \sum_{i=1}^n w_i x_i + w_{n+1} < T_2$$

.....

$$S_m : T_{m-1} \leq \sum_{i=1}^n w_i x_i + w_{n+1} < T_m$$

Öğrenme Kuralı: $\Delta w = c \frac{1}{2} [y_d - y] x \quad \Delta w = -(y_d - y) \phi'(v) x$