Lojistik Regresyon

Giriş

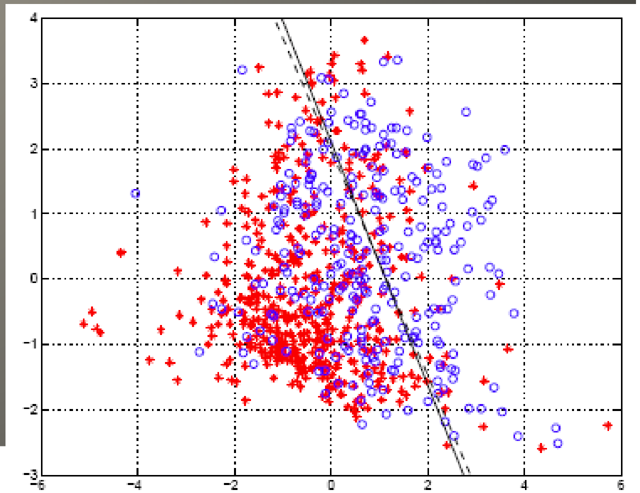
Doç. Dr. İlhan AYDIN

Lojistik Regresyon

- Olasılığa Dayalı Sınıflandırma
  - Sadece sınıfı tahmin etmek yerine, örneğin o sınıf olma olasılığını verin
    - Örneğin, p(y | x)
  - Algılayıcı ile karşılaştırma
    - Algılayıcı bir olasılık tahmini üretmiyor
    - Algılayıcı (ve diğer ayrımcı sınıflandırıcılar) yalnızca ayrımcı bir model üretmekle ilgilenirler.
  - Hatırlatma:

$$0 \leq p(\text{event}) \leq 1$$

$$p(\text{event}) + p(\neg\text{event}) = 1$$



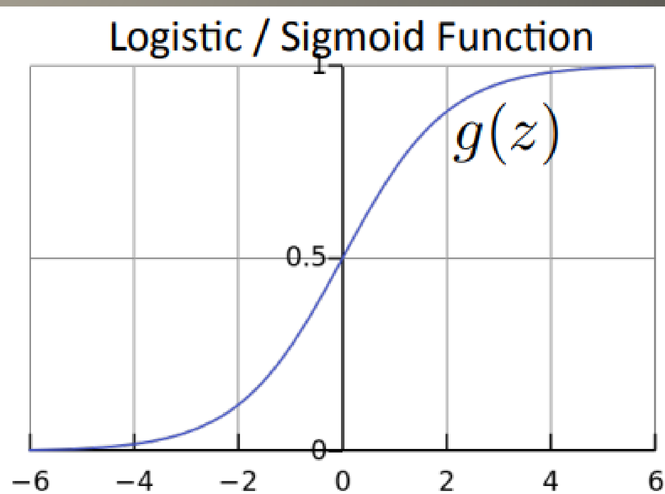
Lojistik Regresyon

- Lojistik Regresyon
  - Ayırt edici işlevleri öğrenmek için olasılıksal bir yaklaşım benimser (yani, bir sınıflandırıcı)
  - $h_0(x)$ $p(y = 1 \mid x; \theta)$'yi vermeli
  - $0 \leq h_0(x) \leq 1$ (Sadece bir eşik ile doğrusal regresyon kullanamazsınız)
  - Lojistik regresyon modeli:

$$h_{\boldsymbol{\theta}}(\boldsymbol{x}) = g\left(\boldsymbol{\theta}^{\mathsf{T}}\boldsymbol{x}\right)$$

$$g(z) = \frac{1}{1 + e^{-z}}$$

$$\boxed{h_{\boldsymbol{\theta}}(\boldsymbol{x}) = \frac{1}{1 + e^{-\boldsymbol{\theta}^{\mathsf{T}}\boldsymbol{x}}}}$$



Lojistik Regresyon

- Hipotez Çıktısının Yorumlanması
  - $h_0(x)$ tahmini $p(y = 1 \mid x; \theta)$

Example: Cancer diagnosis from tumor size

$$\boldsymbol{x} = \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} = \begin{bmatrix} 1 \\ \text{tumorSize} \end{bmatrix}$$

$$h_{\boldsymbol{\theta}}(\boldsymbol{x}) = 0.7$$

→ Tell patient that 70% chance of tumor being malignant

  - Not: $p(y = 0 \mid x; \theta) + p(y = 1 \mid x; \theta) = 1$
  - Öyleyse, $p(y = 0 \mid x; \theta) = 1 - p(y = 1 \mid x; \theta)$

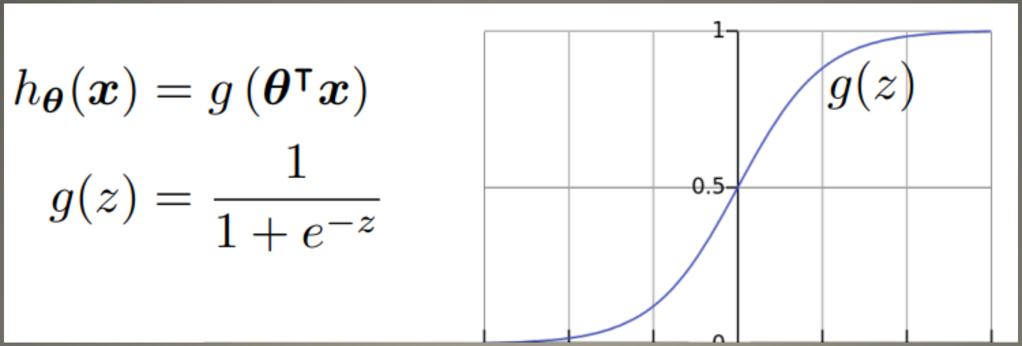
Lojistik Regresyon

- Başka Bir Yorum
  - Eşdeğer olarak, lojistik regresyon şunu varsayar:

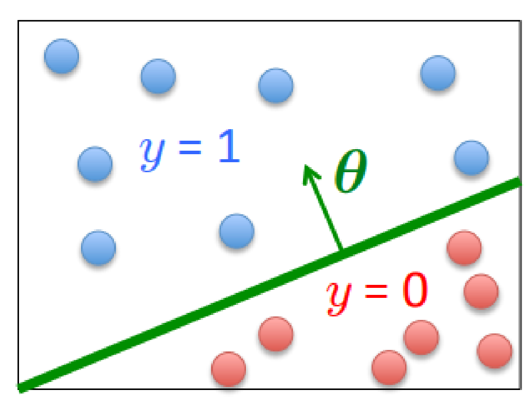
$$\log \underbrace{\boxed{\frac{p(y=1 \mid \boldsymbol{x};\boldsymbol{\theta})}{p(y=0 \mid \boldsymbol{x};\boldsymbol{\theta})}}}_{\text{odds of } y = 1} = \theta_0 + \theta_1 x_1 + \ldots + \theta_d x_d$$

  - Not: Bir olayın lehine olan oran, p / (1 - p) miktarıdır, burada p, olayın olasılığıdır
  - Diğer bir deyişle, lojistik regresyon, log oranlarının x'in doğrusal bir fonksiyonu olduğunu varsayar.

- Lojistik Regresyon



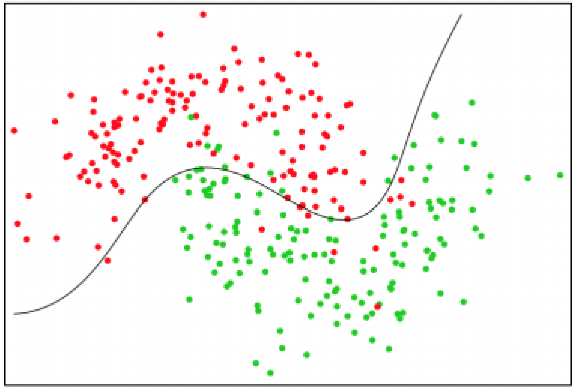
  - $\theta^T x$, negatif örnekler için büyük negatif değerler olmalıdır.
  - $\theta^T x$, pozitif örnekler için büyük pozitif değerler olmalıdır.
- Bir eşik varsayalım ve..
  - $h_0(x)$ >= 0.5 ise y=1 tahmin et
  - $h_0(x)$ < 0.5 ise y=0 tahmin et



Lojistik Regresyon

- Doğrusal Olmayan Karar Sınırı
  - Lineer regresyonda olduğu gibi, temel fonksiyon genişletmesini özelliklere uygulayabilir

$$\boldsymbol{x} = \begin{bmatrix} 1 \\ x_1 \\ x_2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_1 x_2 \\ x_1^2 \\ x_2^2 \\ x_1^2 x_2 \\ x_1 x_2^2 \\ \vdots \end{bmatrix}$$



Lojistik Regresyon

- Lojistik Regresyon
  - $\{(x^1, y^1), (x^2, y^2), .. (x^n, y^n)\}$ verildiğinde $x^i \in \mathbb{R}$, $y^i \in \{0,1\}$
  - Model:

$$h_{\boldsymbol{\theta}}(\boldsymbol{x}) = g\left(\boldsymbol{\theta}^\intercal \boldsymbol{x}\right)$$

$$g(z) = \frac{1}{1 + e^{-z}}$$

$$\boldsymbol{\theta} = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_d \end{bmatrix} \qquad \boldsymbol{x}^\intercal = \begin{bmatrix} 1 & x_1 & \ldots & x_d \end{bmatrix}$$

Lojistik Regresyon

- Lojistik Regresyon Amaç Fonksiyonu
  - Doğrusal regresyonda olduğu gibi sadece kare kaybı kullanamazsınız:

$$J(\boldsymbol{\theta}) = \frac{1}{2n} \sum_{i=1}^{n} \left(h_{\boldsymbol{\theta}} \left(\boldsymbol{x}^{(i)} \right) - y^{(i)} \right)^2$$

  - Lojistik regresyon modelini kullanma

$$h_{\boldsymbol{\theta}}(\boldsymbol{x}) = \frac{1}{1 + e^{-\boldsymbol{\theta}^\top \boldsymbol{x}}}$$

  - Dışbükey olmayan bir optimizasyonla sonuçlanır

Lojistik Regresyon

- Maksimum Olabilirlik Tahmini ile Maliyet Fonksiyonunun Türetilmesi
  - Veri olasılığı şu şekilde verilir: $l(\boldsymbol{\theta}) = \prod_{i=1}^{n} p(y^{(i)} \mid \boldsymbol{x}^{(i)}; \boldsymbol{\theta})$
  - Bu nedenle, olasılığı en üst düzeye çıkaran θ'yi arayın.

$$\boldsymbol{\theta}_{\text{MLE}} = \arg\max_{\boldsymbol{\theta}} l(\boldsymbol{\theta}) = \arg\max_{\boldsymbol{\theta}} \prod_{i=1}^{n} p(y^{(i)} \mid \boldsymbol{x}^{(i)}; \boldsymbol{\theta})$$

  - Çözümü değiştirmeden kayıtları alabilir:

$$\begin{aligned}\boldsymbol{\theta}_{\text{MLE}} &= \arg\max_{\boldsymbol{\theta}} \log \prod_{i=1}^{n} p(y^{(i)} \mid \boldsymbol{x}^{(i)}; \boldsymbol{\theta}) \\ &= \arg\max_{\boldsymbol{\theta}} \sum_{i=1}^{n} \log p(y^{(i)} \mid \boldsymbol{x}^{(i)}; \boldsymbol{\theta})\end{aligned}$$

Lojistik Regresyon

- Maksimum Olabilirlik Tahmini ile Maliyet Fonksiyonunun Türetilmesi
  - Aşağıdaki gibi genişletin

$$\begin{aligned}\boldsymbol{\theta}_{\mathrm{MLE}} &= \arg\max_{\boldsymbol{\theta}} \sum_{i=1}^{n} \log p(y^{(i)} \mid \boldsymbol{x}^{(i)}; \boldsymbol{\theta}) \\ &= \arg\max_{\boldsymbol{\theta}} \sum_{i=1}^{n} \left[y^{(i)} \log p(y^{(i)}=1 \mid \boldsymbol{x}^{(i)}; \boldsymbol{\theta}) + \left(1 - y^{(i)}\right) \log \left(1 - p(y^{(i)}=1 \mid \boldsymbol{x}^{(i)}; \boldsymbol{\theta})\right)\right]\end{aligned}$$

  - Modelde değiştirin ve verim için negatif alın

Logistic regression objective:

$$\min_{\boldsymbol{\theta}} J(\boldsymbol{\theta})$$

$$J(\boldsymbol{\theta}) = -\sum_{i=1}^{n} \left[y^{(i)} \log h_{\boldsymbol{\theta}}(\boldsymbol{x}^{(i)}) + \left(1 - y^{(i)}\right) \log \left(1 - h_{\boldsymbol{\theta}}(\boldsymbol{x}^{(i)})\right)\right]$$

Lojistik Regresyon

- Hedefin Arkasındaki Sezgi

$$J(\boldsymbol{\theta}) = -\sum_{i=1}^{n}\left[y^{(i)} \log h_{\boldsymbol{\theta}}(\boldsymbol{x}^{(i)}) + \left(1 - y^{(i)}\right) \log \left(1 - h_{\boldsymbol{\theta}}(\boldsymbol{x}^{(i)})\right)\right]$$

  - Tek bir örneğin maliyeti

$$\operatorname{cost}\left(h_{\boldsymbol{\theta}}(\boldsymbol{x}), y\right) = \begin{cases} -\log(h_{\boldsymbol{\theta}}(\boldsymbol{x})) & \text{if } y = 1 \\ -\log(1 - h_{\boldsymbol{\theta}}(\boldsymbol{x})) & \text{if } y = 0 \end{cases}$$

  - Amaç fonksiyonunu şu şekilde yeniden yazabilir:

$$J(\boldsymbol{\theta}) = \sum_{i=1}^{n} \operatorname{cost}\left(h_{\boldsymbol{\theta}}(\boldsymbol{x}^{(i)}), y^{(i)}\right)$$

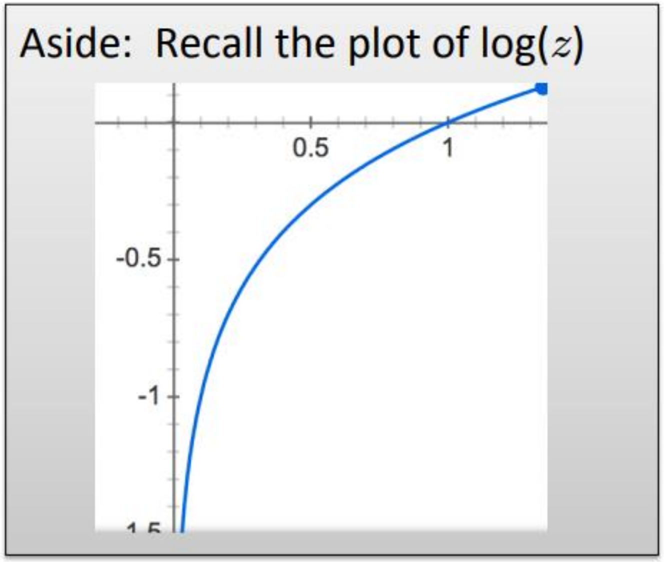
  - Doğrusal regresyonla karşılaştırın:

$$J(\boldsymbol{\theta}) = \frac{1}{2n}\sum_{i=1}^{n}\left(h_{\boldsymbol{\theta}}\left(\boldsymbol{x}^{(i)}\right) - y^{(i)}\right)^2$$

Lojistik Regresyon

- Hedefin Arkasındaki Sezgi

$$\text{cost}\left(h_{\boldsymbol{\theta}}(\boldsymbol{x}), y\right) = \begin{cases} -\log(h_{\boldsymbol{\theta}}(\boldsymbol{x})) & \text{if } y = 1 \\ -\log(1 - h_{\boldsymbol{\theta}}(\boldsymbol{x})) & \text{if } y = 0 \end{cases}$$

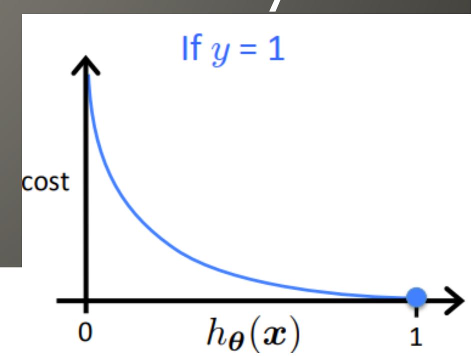


Lojistik Regresyon

- Hedefin Arkasındaki Sezgi

$$\text{cost}\left(h_{\boldsymbol{\theta}}(\boldsymbol{x}), y\right) = \begin{cases} \boxed{-\log(h_{\boldsymbol{\theta}}(\boldsymbol{x})) \quad \text{if } y = 1} \\ -\log(1 - h_{\boldsymbol{\theta}}(\boldsymbol{x})) \quad \text{if } y = 0 \end{cases}$$

  - Eğer y=1 ise,
  - Eğer tahmin doğruysa cost = 0
  - $h_0(x)$->0, cost->∞ olarak
  - Daha büyük hataların daha büyük cezalar alması gerektiğine dair sezgiyi yakalar
  - Örneğin $h_0(x) = 0$ olarak tahmin edin ama y=1

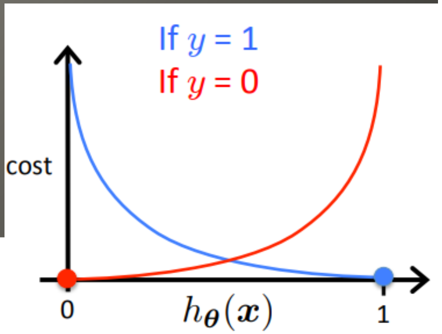


Lojistik Regresyon

- Hedefin Arkasındaki Sezgi

$$\text{cost}\left(h_{\boldsymbol{\theta}}(\boldsymbol{x}), y\right) = \begin{cases} \boxed{-\log(h_{\boldsymbol{\theta}}(\boldsymbol{x})) \quad \text{if } y = 1} \\ -\log(1 - h_{\boldsymbol{\theta}}(\boldsymbol{x})) \quad \text{if } y = 0 \end{cases}$$

  - Eğer y=1 ise,
  - Eğer tahmin doğruysa cost = 0
  - $h_0(x)$->0, cost->∞ olarak
  - Daha büyük hataların daha büyük cezalar alması gerektiğine dair sezgiyi yakalar
  - Örneğin $h_0(x) = 0$ olarak tahmin edin ama y=1



Lojistik Regresyon

- Düzenli Lojistik Regresyon

$$J(\boldsymbol{\theta}) = -\sum_{i=1}^{n}\left[y^{(i)} \log h_{\boldsymbol{\theta}}(\boldsymbol{x}^{(i)}) + \left(1 - y^{(i)}\right) \log\left(1 - h_{\boldsymbol{\theta}}(\boldsymbol{x}^{(i)})\right)\right]$$

  - Lojistik regresyonu tam olarak daha önce olduğu gibi düzenleyebiliriz:

$$\begin{aligned} J_{\text{regularized}}(\boldsymbol{\theta}) &= J(\boldsymbol{\theta}) + \frac{\lambda}{2}\sum_{j=1}^{d}\theta_j^2 \\ &= J(\boldsymbol{\theta}) + \frac{\lambda}{2}\|\boldsymbol{\theta}_{[1:d]}\|_2^2 \end{aligned}$$

Lojistik Regresyon

- Lojistik Regresyon için Gradyan İniş (Gradient Descent)

$$J_{\text{reg}}(\boldsymbol{\theta}) = -\sum_{i=1}^{n}\left[y^{(i)}\log h_{\boldsymbol{\theta}}(\boldsymbol{x}^{(i)}) + \left(1-y^{(i)}\right)\log\left(1-h_{\boldsymbol{\theta}}(\boldsymbol{x}^{(i)})\right)\right] + \frac{\lambda}{2}\|\boldsymbol{\theta}_{[1:d]}\|_2^2$$

  - $\min_{\theta} J(\theta)$ iste
  - θ'yı başlat
  - Yakınsama kadar tekrarlayın

$$\theta_j \leftarrow \theta_j - \alpha\frac{\partial}{\partial\theta_j}J(\boldsymbol{\theta})$$

j=0..d için eşzamanlı güncelleme

  - $h_0(x)$'deki exp() ile iptal etmek için (ln = $\log_e$) doğal logaritmasını kullanın

Lojistik Regresyon

- Lojistik Regresyon için Gradyan İniş (Gradient Descent)

$$J_{\text{reg}}(\boldsymbol{\theta}) = -\sum_{i=1}^{n} \left[y^{(i)} \log h_{\boldsymbol{\theta}}(\boldsymbol{x}^{(i)}) + \left(1 - y^{(i)}\right) \log\left(1 - h_{\boldsymbol{\theta}}(\boldsymbol{x}^{(i)})\right) \right] + \frac{\lambda}{2} \|\boldsymbol{\theta}_{[1:d]}\|_2^2$$

  - $\min_{\theta} J(\theta)$ iste
  - θ'yı başlat
  - Yakınsama kadar tekrarlayın

j=0..d için eşzamanlı güncelleme

$$\theta_0 \leftarrow \theta_0 - \alpha \sum_{i=1}^{n} \left(h_{\boldsymbol{\theta}}\left(\boldsymbol{x}^{(i)}\right) - y^{(i)} \right)$$

$$\theta_j \leftarrow \theta_j - \alpha \left[\sum_{i=1}^{n} \left(h_{\boldsymbol{\theta}}\left(\boldsymbol{x}^{(i)}\right) - y^{(i)} \right) x_j^{(i)} + \lambda \theta_j \right]$$

Lojistik Regresyon

- Lojistik Regresyon için Gradyan İniş (Gradient Descent)
  - θ'yı başlat
  - Yakınsama kadar tekrarlayın

$$\theta_0 \leftarrow \theta_0 - \alpha \sum_{i=1}^{n} \left(h_{\boldsymbol{\theta}} \left(\boldsymbol{x}^{(i)} \right) - y^{(i)} \right)$$

$$\theta_j \leftarrow \theta_j - \alpha \left[\sum_{i=1}^{n} \left(h_{\boldsymbol{\theta}} \left(\boldsymbol{x}^{(i)} \right) - y^{(i)} \right) x_j^{(i)} + \lambda \theta_j \right]$$

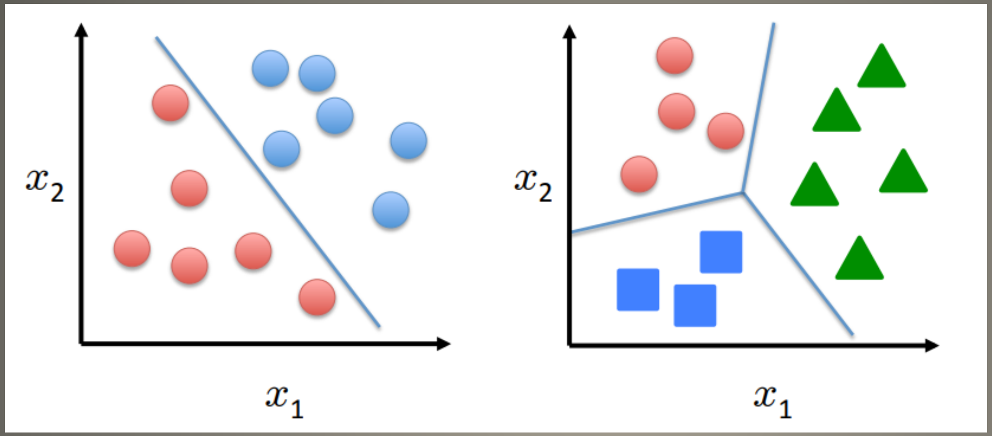
j=0..d için eşzamanlı güncelleme

  - Bu, doğrusal regresyonla AYNI görünüyor !!!
  - 1/n sabitini yok say
  - Ancak, modelin formu çok farklıdır:

$$h_{\boldsymbol{\theta}}(\boldsymbol{x}) = \frac{1}{1 + e^{-\boldsymbol{\theta}^{\mathsf{T}} \boldsymbol{x}}}$$

Lojistik Regresyon

- Çok Sınıflı Sınıflandırma
    - İkili Çoklu



- Hastalık teşhisi: sağlıklı / soğuk algınlığı / grip / pnömoni
- Nesne sınıflandırması: masa / sandalye / monitör / kitaplık

Lojistik Regresyon

- Çok Sınıflı Sınıflandırma
  - 2 sınıf için:

$$h_{\boldsymbol{\theta}}(\boldsymbol{x}) = \frac{1}{1 + \exp(-\boldsymbol{\theta}^{\top}\boldsymbol{x})} = \frac{\exp(\boldsymbol{\theta}^{\top}\boldsymbol{x})}{\boxed{1} + \boxed{\exp(\boldsymbol{\theta}^{\top}\boldsymbol{x})}}$$

weight assigned to $y = 0$

weight assigned to $y = 1$

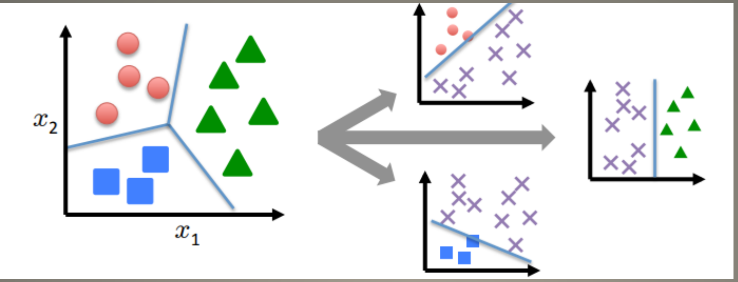
  - {1, ..., C} C sınıfları için:

$$p(y = c \mid \boldsymbol{x}; \boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_C) = \frac{\exp(\boldsymbol{\theta}_c^{\top}\boldsymbol{x})}{\sum_{c=1}^{C} \exp(\boldsymbol{\theta}_c^{\top}\boldsymbol{x})}$$

  - Softmax denir

Lojistik Regresyon

- Çok Sınıflı Sınıflandırma
  - Bire karşı diğerlerini bölün



  - y = i olasılığını tahmin etmek için her i sınıfı için bir lojistik regresyon sınıflandırıcı eğitin.

$$h_c(\boldsymbol{x}) = \frac{\exp(\boldsymbol{\theta}_c^\intercal \boldsymbol{x})}{\sum_{c=1}^{C} \exp(\boldsymbol{\theta}_c^\intercal \boldsymbol{x})}$$

Lojistik Regresyon

- Çok Sınıflı Lojistik Regresyon Uygulaması
  - c sınıfı için model olarak:

$$h_c(\boldsymbol{x}) = \frac{\exp(\boldsymbol{\theta}_c^{\mathsf{T}}\boldsymbol{x})}{\sum_{c=1}^{C}\exp(\boldsymbol{\theta}_c^{\mathsf{T}}\boldsymbol{x})}$$ kullanın.

    - Gradyan iniş, tüm modeller için tüm parametreleri aynı anda günceller.
      - Öncekiyle aynı türev, sadece yukarıdaki $h_c(x)$ ile.
    - Sınıf etiketini en olası etiket olarak tahmin edin.

$$\max_{c} h_c(\boldsymbol{x})$$

Lojistik Regresyon