MAXIMUM LIKELIHOOD ESTIMATION

• Verilen örnekler (x_i, y_i) ∈ R p × {0, 1}, i = 1,. . . n, p (x_i) = P (y_i = 1 | x_i) olsun ve

$$\log\left(\frac{p(x_i)}{1-p(x_i)}\right) = \beta^T x_i, \quad i = 1, \ldots n$$

• Katsayıların bir tahminini β̂ oluşturmak için, maksimum olasılık ilkesini kullanacağız. Yani, örneklerin bağımsızlığını varsayarsak, olasılık (x_i şartına bağlı, i = 1,.. N)

$$\begin{aligned} L(\beta) &= \prod_{i\,:\,y_i=1} p(x_i) \cdot \prod_{i\,:\,y_i=0} (1-p(x_i)) \\ &= \prod_{i=1}^{n} p(x_i)^{y_i} \big(1-p(x_i)\big)^{1-y_i}. \end{aligned}$$

Bu olasılık kriterini maksimize etmek için β̂ seçeceğiz.

• Bir işlevi maksimize etmenin, bir işlevin günlüğünü maksimize etmekle aynı olduğunu unutmayın (çünkü log monoton artmaktadır). Bu nedenle β̂, log olasılığını en üst düzeye çıkarmak için eşit olarak seçilir.

$$\ell(\beta) = \sum_{i=1}^{n} y_i \log p(x_i) + (1-y_i)\log\big(1-p(x_i)\big).$$

Bunu şu şekilde yeniden düzenlemeye yardımcı olur:

$$\begin{aligned} \ell(\beta) &= \sum_{i=1}^{n} y_i \big[\log p(x_i) - \log\big(1-p(x_i)\big)\big] + \log\big(1-p(x_i)\big) \\ &= \sum_{i=1}^{n} y_i \log\left(\frac{p(x_i)}{1-p(x_i)}\right) + \log\big(1-p(x_i)\big). \end{aligned}$$

Son olarak, log(p(x_i)/(1 − p(x_i))) = x^T_i β için ve 1 − p(x_i) = 1/(1 + exp(x^T_i β)), i = 1, … n,

$$\ell(\beta) = \sum_{i=1}^{n} y_i(x_i^T \beta) - \log\left(1 + \exp(x_i^T \beta)\right). \qquad (1)$$

Regresyon için en küçük kareler kriterinden farklı olarak, bu kriter ℓ (β) 'nin maksimize edicisi için kapalı form ifadesine sahip olmadığını görebilirsiniz (örneğin, kısmi türevlerini alıp sıfıra eşitlemeyi deneyin). Dolayısıyla β̂'yi bulmak için bir optimizasyon algoritması çalıştırmalıyız.

- Biraz dikkat çekici bir şekilde, tekrarlanan ağırlıklı en küçük kareler regresyonlarını çalıştırarak (1) maksimize edebiliriz! Biraz optimizasyon öğrenmiş olanlarınız için, bu aslında Newton'un yönteminin bir somutlaştırmasıdır. Kriter (1) 'e uygulandığında, bunu yinelemeli olarak yeniden ağırlıklandırılmış en küçük kareler veya IRLS olarak adlandırıyoruz.

- Kısaca: lojistik regresyonda β̂ tahmini, lineer regresyonda olduğundan daha fazla yer alır, ancak bunu yinelemeli olarak lineer regresyon yazılımı kullanarak yapmak mümkündür.