Computer Vision I

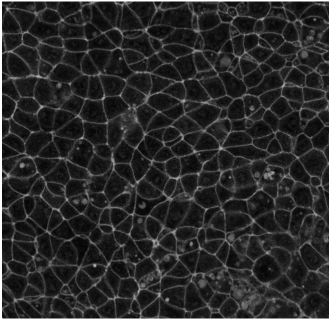
Bjoern Andres, Holger Heidrich, Jannik Presberger

Machine Learning for Computer Vision
TU Dresden

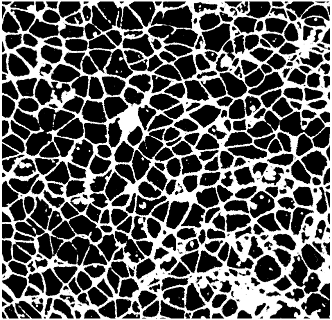
Winter Term 2023/2024

Pixel classification

Digital image[1] $f\colon V \to C$



Classification $y\colon V \to \{0,1\}$



[1]By courtesy of Stephan Grill and his lab at the MPI of Molecular Cell Biology and Genetics.

Suppose we can construct a function $c\colon V \to \mathbb{R}$ wrt. a ditial image $f\colon V \to C$ in such a way that for any pixel $v \in V$:

- $c_v < 0$ if we consider $y_v = 1$ to be the right decision
- $c_v > 0$ if we consider $y_v = 0$ to be the right decision.

Definition 1. For any set V of pixels and any function $c\colon V \to \mathbb{R}$, the instance of the **trivial pixel classification problem** wrt. c has the form

$$\min_{y \in \{0,1\}^V} \sum_{v \in V} c_v \, y_v \tag{1}$$

In case the decision y_v for a pixel v depends on the color $f(v)$ of that pixel only, we can in principle

- construct a function $\xi\colon C \to \mathbb{R}$
- define $c_v = \xi(f(v))$ for any $v \in V$.

In practice, this task is supported by carefully designed GUIs.

In case the decision y_v for a pixel v depends on the colors of all pixels in a neighborhood $N(v) \subseteq V$ around v, we can in principle

- construct, for any pixel v, a function $\xi_v\colon C^{N(v)} \to \mathbb{R}$ that assigns a real number $\xi_v(f')$ to any coloring $f'\colon N(v) \to C$ of the neighborhood $N(v)$ of v
- define $c_v = \xi(f_{N(v)})$ for any $v \in V$.

In practice, this task is typically addressed by **machine learning**.

Random variables:

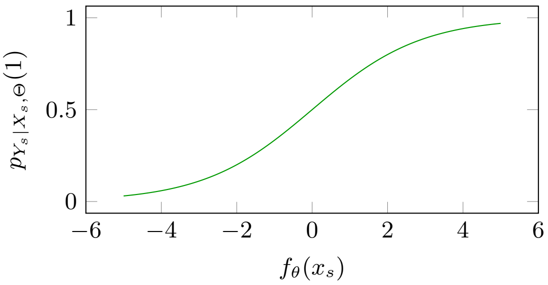
- For any sample $s \in S$, let X_s be a random variable whose value is a vector $x_s \in \mathbb{R}^V$, the **attribute vector** of s
- For any sample $s \in S$, let Y_s be a random variable whose value is a binary number $y_s \in \{0, 1\}$, the **label** of s
- For any $v \in V$, let Θ_v be a random variable whose value is a real number $\theta_v \in \mathbb{R}$, a **parameter** of the linear function we seek to learn

Probabilistic model:

$$P(X, Y, \Theta) = \prod_{s \in S} (P(Y_s \mid X_s, \Theta) P(X_s)) \prod_{v \in V} P(\Theta_v) \tag{2}$$

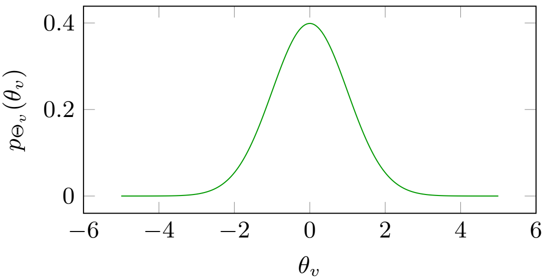
- **Logistic distribution**

$$\forall s \in S : \qquad p_{Y_s \mid X_s, \Theta}(1) = \frac{1}{1 + 2^{-f_\theta(x_s)}} \tag{3}$$



- **Normal distribution** with $\sigma \in \mathbb{R}^+$:

$$\forall v \in V : \qquad p_{\Theta_v}(\theta_v) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\theta_v^2/2\sigma^2} \tag{4}$$



The learning problem consists in maximizing the probability

$$\begin{aligned} P(\Theta \mid X, Y) &= \frac{P(X, Y, \Theta)}{P(X, Y)} \\ &= \frac{P(Y \mid X, \Theta)\, P(X)\, P(\Theta)}{P(X, Y)} \\ &\propto P(Y \mid X, \Theta)\, P(\Theta) \\ &= \prod_{s \in S} P(Y_s \mid X_s, \Theta) \prod_{v \in V} P(\Theta_v) \end{aligned}$$

The inference problem consists in maximizing the probability

$$P(Y \mid X, \Theta) = \prod_{s \in S} P(Y_s \mid X_s, \Theta)$$

Lemma. Estimating maximally probable parameters θ, given attributes x and labels y, i.e.,

$$\underset{\theta \in \mathbb{R}^m}{\operatorname{argmax}} \quad p_{\Theta|X,Y}(\theta, x, y)$$

is equivalent ot the optimization problem

$$\min_{\theta \in \Theta} \quad \lambda R(\theta) + \frac{1}{|S|} \sum_{s \in S} L(f_\theta(x_s), y_s) \tag{5}$$

with L, R and λ such that

$$\forall r \in \mathbb{R}\ \forall \hat{y} \in \{0,1\}\colon \quad L(r, \hat{y}) = -\hat{y}r + \log\left(1 + 2^r\right) \tag{6}$$

$$\forall \theta \in \Theta\colon \quad R(\theta) = \|\theta\|_2^2 \tag{7}$$

$$\lambda = \frac{\log e}{2\sigma^2} \quad . \tag{8}$$

It is called the l_2-regularized **logistic regression problem** with respect to x, y and σ.

Proof. Firstly,

$$\begin{aligned}
&\underset{\theta \in \mathbb{R}^m}{\operatorname{argmax}} \quad p_{\Theta|X,Y}(\theta, x, y) \\
= &\underset{\theta \in \mathbb{R}^m}{\operatorname{argmax}} \quad \prod_{s \in S} p_{Y_s|X_s,\Theta}(y_s, x_s, \theta) \prod_{v \in V} p_{\Theta_v}(\theta_v) \\
= &\underset{\theta \in \mathbb{R}^m}{\operatorname{argmax}} \quad \sum_{s \in S} \log p_{Y_s|X_s,\Theta}(y_s, x_s, \theta) + \sum_{v \in V} \log p_{\Theta_v}(\theta_v) \qquad (9)
\end{aligned}$$

Secondly,

$$\begin{aligned}
&\log p_{Y_s|X_s,\Theta}(y_s, x_s, \theta) \\
= \quad &y_s \log p_{Y_s|X_s,\Theta}(1, x_s, \theta) + (1 - y_s) \log p_{Y_s|X_s,\Theta}(0, x_s, \theta) \\
= \quad &y_s \log \frac{p_{Y_s|X_s,\Theta}(1, x_s, \theta)}{p_{Y_s|X_s,\Theta}(0, x_s, \theta)} + \log p_{Y_s|X_s,\Theta}(0, x_s, \theta) \qquad (10)
\end{aligned}$$

Thus, with (3) and (4):

$$\underset{\theta \in \mathbb{R}^m}{\operatorname{argmin}} \quad \sum_{s \in S} \left(-y_s \langle \theta, x_s \rangle + \log \left(1 + 2^{\langle \theta, x_s \rangle} \right) \right) + \frac{\log e}{2\sigma^2} \|\theta\|_2^2 \qquad (11)$$

Lemma 1. The objective function

$$\varphi(\theta) \;=\; \lambda R(\theta) + \frac{1}{|S|}\sum_{s \in S} L(f_\theta(x_s), y_s) \qquad (12)$$

of the l_2-regularized logistic regression problem is convex.

The problem can be solved, e.g., by the steepest descent algorithm with a tolerance parameter $\epsilon \in \mathbb{R}_0^+$:

$\theta := 0$
repeat
    $d := \nabla \varphi(\theta)$
    $\eta := \operatorname{argmin}_{\eta' \in \mathbb{R}} \varphi(\theta - \eta' d)$ (line search)
    $\theta := \theta - \eta d$
    if $\|d\| < \epsilon$
        return θ

Lemma: Estimating maximally probable labels y, given attributes x' and parameters θ, i.e.,

$$\operatorname*{argmax}_{y \in \{0,1\}^S} \quad p_{Y|X,\Theta}(y, x', \theta) \tag{13}$$

is equivalent to the inference problem

$$\min_{y' \in \{0,1\}^S} \sum_{s \in S} L(f_\theta(x_s), y'_s) \ . \tag{14}$$

It has the solution

$$\forall s \in S' : \quad y_s = \begin{cases} 1 & \text{if } f_\theta(x'_s) > 0 \\ 0 & \text{otherwise} \end{cases} \ . \tag{15}$$

Proof. Firstly,

$$\begin{aligned}
&\underset{y \in \{0,1\}^{S'}}{\operatorname{argmax}} \quad p_{Y|X,\Theta}(y, x', \theta) \\
=\ &\underset{y \in \{0,1\}^{S'}}{\operatorname{argmax}} \quad \prod_{s \in S'} p_{Y_s|X_s,\Theta}(y_s, x'_s, \theta) \\
=\ &\underset{y \in \{0,1\}^{S'}}{\operatorname{argmax}} \quad \sum_{s \in S'} \log p_{Y_s|X_s,\Theta}(y_s, x'_s, \theta) \\
=\ &\underset{y \in \{0,1\}^{S'}}{\operatorname{argmax}} \quad \sum_{s \in S'} \left(y_s \log \frac{p_{Y_s|X_s,\Theta}(1, x'_s, \theta)}{p_{Y_s|X_s,\Theta}(0, x'_s, \theta)} + \log p_{Y_s|X_s,\Theta}(0, x'_s, \theta) \right) \\
=\ &\underset{y \in \{0,1\}^{S'}}{\operatorname{argmin}} \quad \sum_{s \in S'} \left(-y_s f_\theta(x'_s) + \log\left(1 + 2^{f_\theta(x'_s)}\right)\right) \\
=\ &\underset{y \in \{0,1\}^{S'}}{\operatorname{argmin}} \quad \sum_{s \in S'} L(f_\theta(x'_s), y_s) \enspace .
\end{aligned}$$

Secondly,

$$\min_{y \in \{0,1\}^{S'}} \sum_{s \in S'} \left(-y_s f_\theta(x'_s) + \log\left(1 + 2^{f_\theta(x'_s)}\right)\right) = \sum_{s \in S'} \max_{y_s \in \{0,1\}} y_s f_\theta(x'_s) \enspace .$$