Tutorial 7, Mar 8, 2024

Linear Regression

- Given data samples $(x_1, y_1), \ldots, (x_n, y_n)$, we want to find a linear model with parameters a, b such that $ax + b$ best fits the data
- Use the squared error, so $(a^*, b^*) = \underset{(a,b)}{\operatorname{argmin}} \sum_{i=1}^{n} (y_i - (ax_i + b))^2$
- Let $\boldsymbol{Y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$, $\boldsymbol{H} = \begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix}$, $\boldsymbol{v} = \begin{bmatrix} a \\ b \end{bmatrix}$, $\boldsymbol{Z} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}$ then $\boldsymbol{Y} = \boldsymbol{H}\boldsymbol{v} + \boldsymbol{Z}$
- The optimization becomes $\boldsymbol{v}^* = \underset{\boldsymbol{v}}{\operatorname{argmin}} \|\boldsymbol{Y} - \boldsymbol{H}\boldsymbol{v}\|^2$
  - The closed-form solution is $\boldsymbol{v}^*(\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^T\boldsymbol{Y}$
- If we assume that the noise $\boldsymbol{Z}$ is Gaussian with zero mean, then the least-squares solution is equivalent to the ML estimate of $\boldsymbol{Y}$ given $\boldsymbol{X}$
- What if $\boldsymbol{Z}$ is exponentially distributed according to ce^{-z}?
  - Now all the Z_i are positive – this means that instead of having the line of best fit passing through the data, it should now be completely below the data, since all the noise is positive
  - Optimize with the condition $y_i - (ax_i + b) \geq 0$ for all i
  - $\boldsymbol{v}^* = \underset{\boldsymbol{v}}{\operatorname{argmax}} \prod_{i=1}^{n} f_z(y_i - (ax_i + b)) = \underset{\boldsymbol{v}}{\operatorname{argmax}} \exp\left(-\sum_{i=1}^{N} (y_i - (ax_i + b))\right)$
  - NLL: $\underset{\boldsymbol{v}}{\operatorname{argmax}} \sum_{i=1}^{N} (y_i - (ax_i + b))$
  - This has no closed-form solution, but it is a linear program that can be solved quickly using numerical methods