Artem Maevskiy
amaevskij@hse.ru

Machine Learning and Data Mining
Assignment 3

2019/11/27

Overview

The below assignment consists of theoretical and practical parts. Please prepare the solutions for the theoretical part as PDF scans of your derivations on paper (alternatively you can use any software, e.g. LaTeX, to generate PDF with formulas). For the practical part, please prepare the solution in jupyter notebook format. The solutions should be clear and easy to follow.

Please send your PDF and .ipynb files to amaevskij@hse.ru under the following subject:
MLDM-2019-HW3-<YOUR LASTNAME>.

The deadline is: **23:59, 18.12.2019**.

1 Theory

1.1 Bayesian linear regression (0.5 points)

Consider a Bayesian linear regression model: $P(\mathbf{y}\,|\,\mathbf{X},\mathbf{w}) = \mathcal{N}(\mathbf{X}\mathbf{w}, \mathbb{I}\sigma_\epsilon^2)$, where $\mathbf{y}$ is a column vector of targets, $\mathbf{X}$ — design matrix (each row is a vector of features of a given object), $\mathbf{w}$ — column vector of weights, $\mathbb{I}$ — identity matrix, σ_ϵ — fixed parameter; with a normal prior on the weights: $P(\mathbf{w}) = \mathcal{N}(\mathbf{0}, \Sigma_\mathbf{w})$, where $\Sigma_\mathbf{w}$ is a fixed covariance matrix. Derive analytical formula for the posterior distribution $P(\mathbf{w}\,|\,\mathbf{X},\mathbf{y})$.

2 Practice

2.1 Generating digits (0.5 points + 0.1 points bonus)

Train a GAN or VAE on the MNIST dataset, conditioning on the labels — i.e. your generative model should take the digit you want it to generate as an input. Use one-hot encoding for this input. Provide the following plots along with the code:

1. Learning curves for both train and validation losses.
2. At least 10 generated images per each of the digits from 0 to 9 (i.e. at least 10 zeros, 10 ones and so on). Ideally all 100 images should be placed in a single image in a 10 by 10 grid.
3. *(for the bonus points)* Generate images on inputs interpolated between a pair of digits. I.e. let i_a and i_b be the one-hot encoded input vectors corresponding to digits a and b, respectively. Then, generate images for inputs $\alpha i_a + (1-\alpha) i_b$ for all $\alpha \in \left\{\frac{0}{100}, \frac{1}{100}, \ldots, \frac{99}{100}, \frac{100}{100}\right\}$ for some fixed value of the latent space vector z. Pick any pair of different digits a and b of your choice for this task.