Activities 03

Deep Learning Lab

September 28, 2018

1 Assignment 1

Consider the polynomial p given by

$$p(x) = x^3 + 2x^2 - 4x - 8 = \sum_{i=1}^{4} w_i^* x^{i-1},$$

where $\mathbf{w}^* = [-8, -4, 2, 1]^T$.

Consider also a dataset $\mathcal{D} = \{(x_i, y_i)\}_1^N$, where $y_i = p(x_i) + \epsilon_i$, and each ϵ_i is drawn from a normal distribution with mean zero and standard deviation $\sigma = 1/2$.

If the vector $\mathbf{w}^*$ were unknown, linear regression could estimate it given the dataset $\mathcal{D}$. This would require applying a feature map to transform the original dataset $\mathcal{D}$ into an expanded dataset $\mathcal{D}' = \{(\mathbf{x}_i, y_i)\}_1^N$, where $\mathbf{x}_i = [1, x_i, x_i^2, x_i^3]^T$.

Such data generation and expansion is partially illustrated in the code presented below.

Listing 1: Polynomial regression dataset generation (incomplete).

```python
def create_dataset(w_star, x_range, sample_size, sigma, seed=None):
    random_state = np.random.RandomState(seed)

    x = random_state.uniform(x_range[0], x_range[1], (sample_size))
    X = np.zeros((sample_size, w_star.shape[0]))
    for i in range(sample_size):
        X[i, 0] = 1.
        for j in range(1, w_star.shape[0]):
            X[i, j] = ? # Incomplete

    y = X.dot(w_star)
    if sigma > 0:
        y += random_state.normal(0.0, sigma, sample_size)

    return X, y
```

1. Adapt the snippet presented in Pgs. 32-33 to perform polynomial regression using a dataset $\mathcal{D}'$ created using the code presented above. More specifically, find an estimate of $\mathbf{w}^* = [-8, -4, 2, 1]^T$ supposing that such

vector is unknown. Each x_i should be in the interval $[-3, 2]$. Use a sample of size 100 created with a seed of 0 for training, and a sample of size 100 created with a seed of 1 for validation. Let $\sigma = 1/2$.

2. Find a suitable learning rate and number of iterations for gradient descent.

3. Plot the polynomial defined by $\mathbf{w}^*$ and the polynomial defined by your estimate $\hat{\mathbf{w}}$. Plot the training dataset.

4. Observe what happens when the training dataset is reduced to 50, 10, and 5 observations.

5. Observe what happens when σ is increased to 2, 4, and 8.

Deliver your source code together with a short report documenting your experiments.