COMP9727 Recommender Systems
Tutorial 1

Review of Basic ML concepts and Linear Algebra

Change Log:

Last updated: Sunday 29th May 2022 at 12:37 pm

$$Q3: \quad \sum_{i=1}^{n} |w_i| \Rightarrow \sum_{j=1}^{d} |w_j|$$

Q1 - Matrix

a). Given a matrix A, find the **eigenvalues** and **eigenvectors** of A, A^2, A^{-1} and $A + 4I$.

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

b). Find the **rank**, **trace** and **determinant** of A, A^2, A^{-1} and $A + 4I$.
c). Find A^T, $A^T A$ and AA^T, and their corresponding **eigenvalues** and **eigenvectors**.

Q2 - Gradient Descent

Assume you are training a regression model $y = wx + b$, given a cost function $f(w, b)$ which is also known as Mean Squared Error, find the gradient and its gradient descent rule.

$$f(w, b) = \frac{1}{N} \sum_{i=1}^{N} \big(y_i - (wx_i + b)\big)^2$$

Q3 - Regularization

Overfitting is a common problem in machine learning when training a model. Regularization is a process of introducing additional information in order to prevent overfitting. The common way is adding the L^p norm into the loss function. If a model's parameter w has d components,

$$\|w\|_p = \Big(\sum_{j=1}^{d} |w_j|^p\Big)^{\frac{1}{p}}$$

Find the gradient descent rule for Q2's loss function by adding L1 and L2 regularisation respectively. (Hint: L1 regularization is $\|w\|_1$ and L2 regularization is $\|w\|_2^2$.)

Extra Question: What's the difference between regularization and normalization?

Q4 - Collaborative Filtering: Linear Algebra Approach

In this lecture, we've discussed several collaborative filtering methods. However, how to conduct the collaborative filtering in the situation that the rating matrix is not provided? Alex finds a collaborative filtering problem with 4 users, 4 movies, and $k = 2$ (k is the number of features for user/movie), specified by matrices U(for users) and V(for movies):

$$U = \begin{pmatrix} 1 & 2 \\ 3 & 1 \\ 2 & 5 \\ 1 & 3 \end{pmatrix}, V = \begin{pmatrix} 2 & 3 \\ 1 & 4 \\ 5 & 2 \\ 3 & 1 \end{pmatrix}$$

Assumes that bias are all zero (i.e.,$b_U^j = 0, b_V^i = 0$ for all j and i). Assuming that the first entry has index 1, what is the predicted value for user 3's rating of movie 1?