

Машинное обучение.

Домашнее задание №5

Метод главных компонент работает хорошо лишь при условии, что выборка лежит в линейном подпространстве признакового пространства. Если же выборка лежит в нелинейной многообразии, то найти новое признаковое пространство можно с помощью ядрового перехода в методе главных компонент.

Задача 1. Пусть u_1, \dots, u_d — собственные векторы матрицы $X^T X$, а $\lambda_1, \dots, \lambda_d$ — соответствующие им собственные значения. Выразите ненулевые собственные значения и собственные векторы матрицы $X X^T$ через u_1, \dots, u_d и $\lambda_1, \dots, \lambda_d$. Выразите ненулевые собственные векторы и собственные значения матрицы $X^T X$ через собственные векторы и собственные значения матрицы $X X^T$.

Матрица $X X^T$ — это матрица Грама, которая может быть вычислена и в спрямляющем пространстве с помощью функции ядра: $(X X^T)_{ij} = K(x_i, x_j)$. Для данной матрицы можно найти собственное разложение, а через него выразить собственные значения и собственные векторы ковариационной матрицы $X^T X$. Однако данные выражения не будут вычислимыми, поскольку в них будут входить признаковые описания объектов в спрямляющем пространстве $\varphi(x)$. Тем не менее, нас интересуют лишь проекции $z_i = U^T x_i$.

Задача 2. Покажите, как найти новые признаковые описания объектов $z_i = U^T x_i$, зная лишь собственные векторы v_1, \dots, v_d матрицы $X X^T$ и функцию ядра $K(x, z)$.