

1.

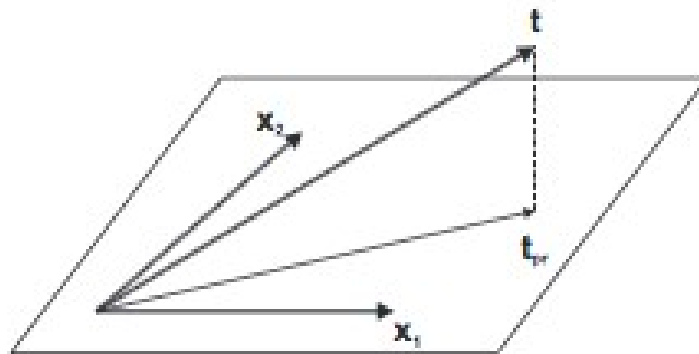
Модель линейной регрессии:

$$t = \sum_{j=1}^d w_j x(j) + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \sigma^2).$$

Поиск весов ω с помощью максимизации правдоподобия выборки в этой модели эквивалентен методу наименьших квадратов:

$$\|t - Xw\|^2 \rightarrow \min_w.$$

Геометрическая интерпретация — поиск проекции вектора t на гиперплоскость с направляющими векторами x_1, \dots, x_d .



2.

`LearningRateScheduler` — встроенная процедура обратного вызова. Она регулирует скорость обучения с течением времени, используя функцию, которую вы уже написали. Эта функция возвращает желаемую скорость обучения. Входной параметр — индекс эпохи.