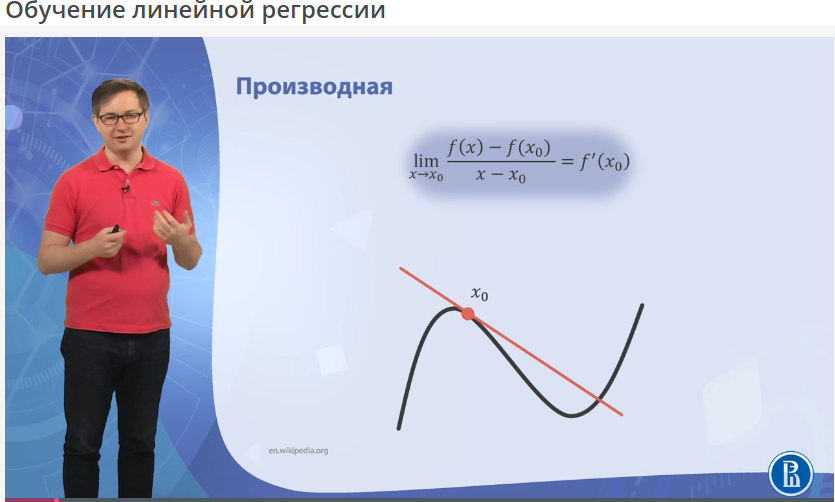


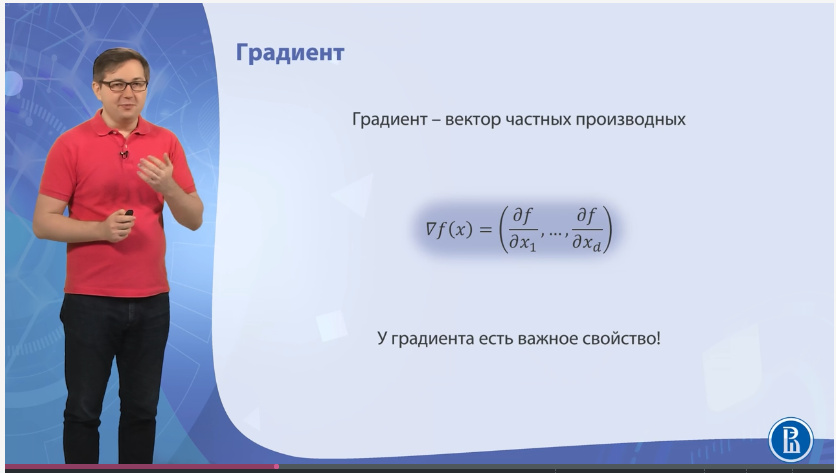
Пока забудем про матричную форму



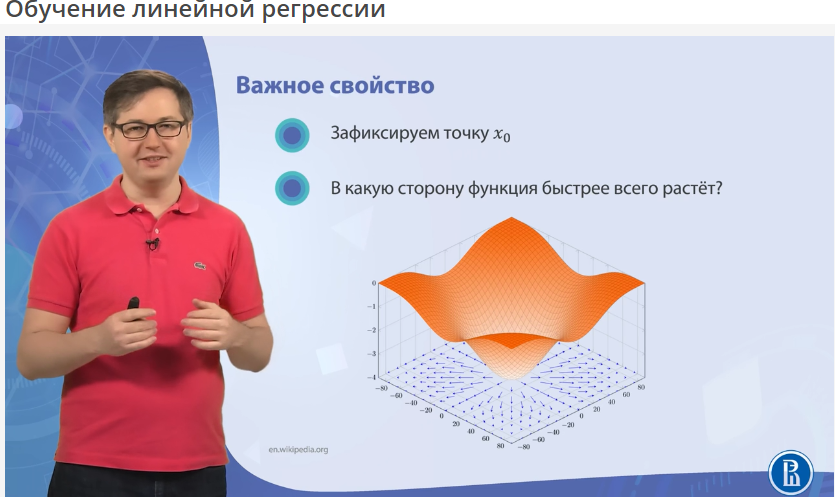
Производная – характеристика скорости роста функции – угол наклона касательной к графику функции в точке x0



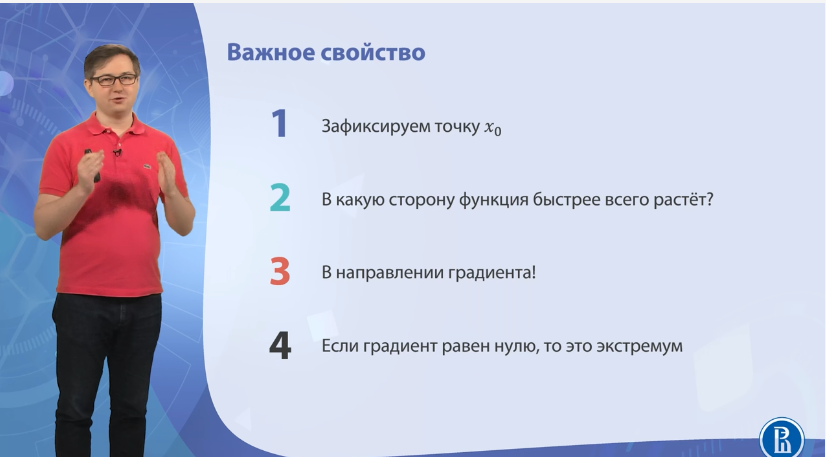
Точки минимума и максимума



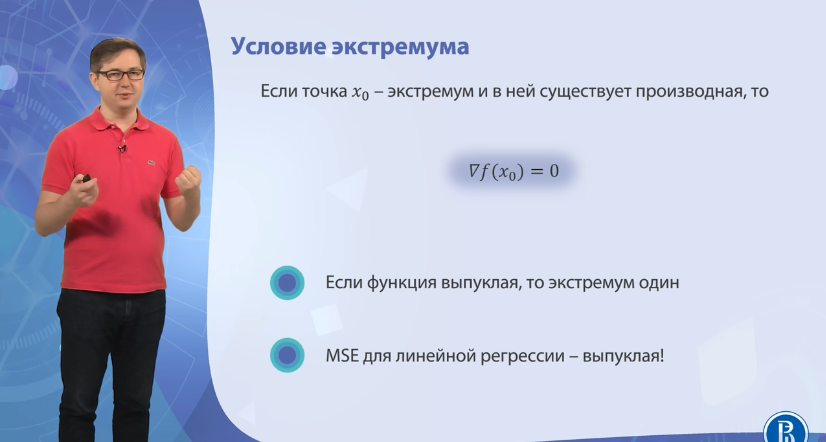
Если функция зависит от нескольких параметров, то градиент



Вектор градиента показывает, в какую сторону функция растет быстрее всего

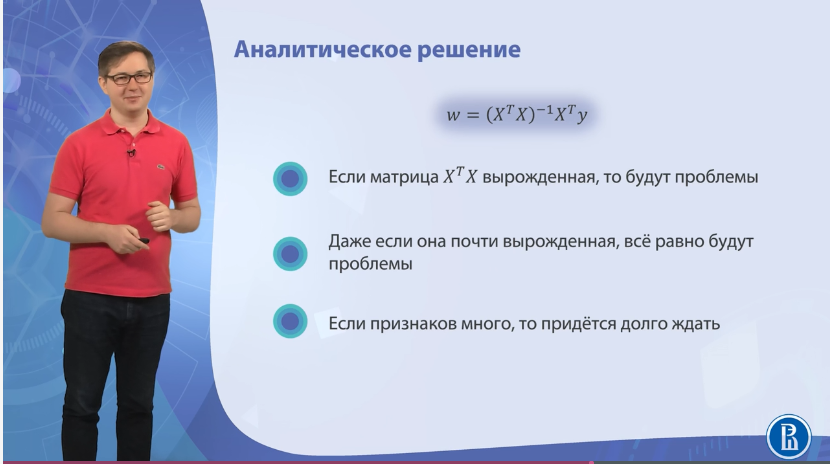


«минус-градиент» показывает, в какую сторону быстрее всего убывает функция





Получается оптимальный вектор w. Произведение матриц XTX – ковариационная матрица

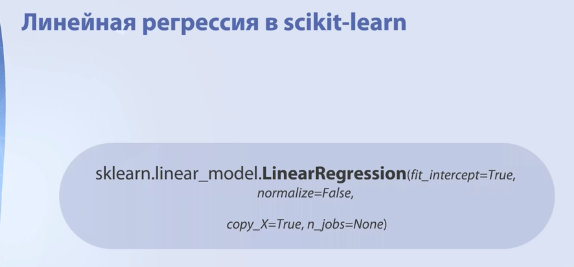


Вы́рожденная ма́трица — квадратная матрица, определитель которой равен нулю, и траспонированной матрицы к ней не существует

Если матрица n\*n, то найти обратную к ней – порядка n3 операций

Формула в целом применима, если выборка небольшая и признаков немного

Если бы функция потерь была бы не MSE, то не факт, что удалось бы найти решение



Fit\_intercept – нужно ли отдельно строить w0, или мы уже включили единичный признак

