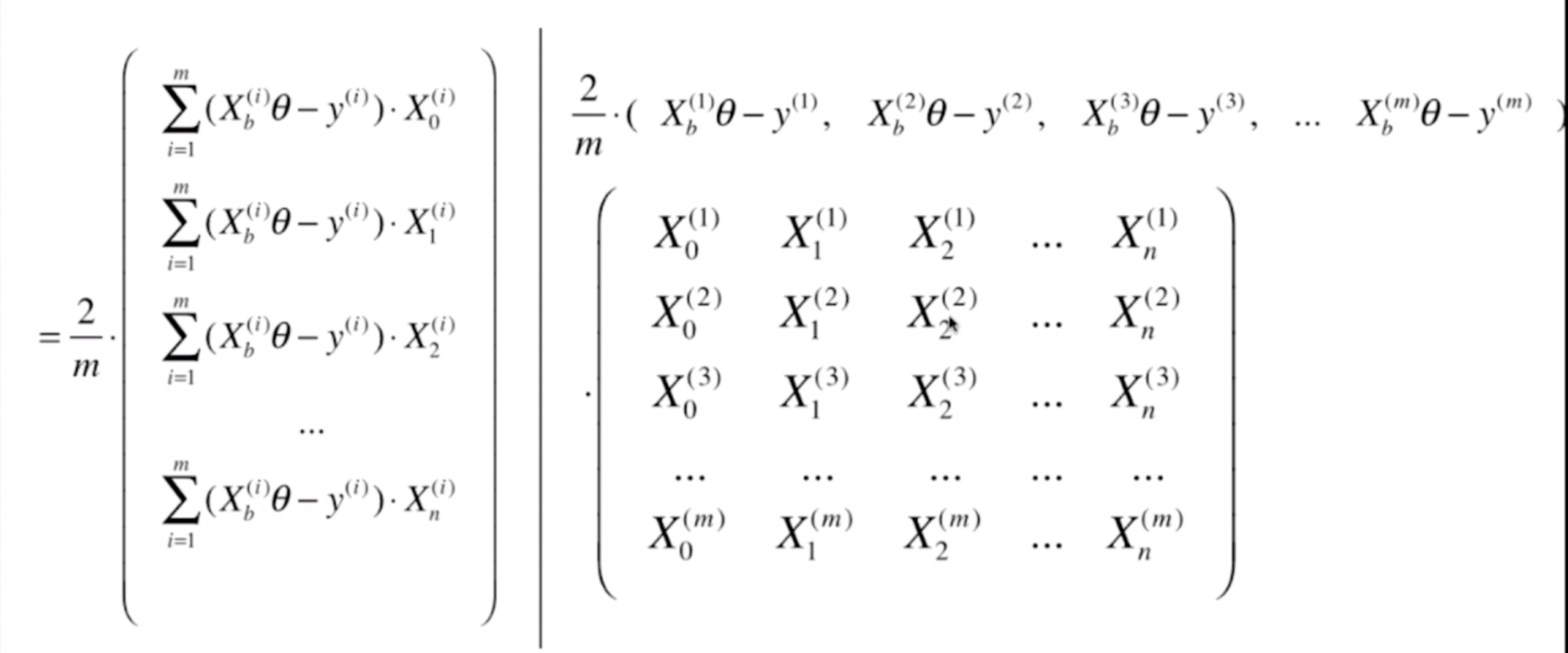


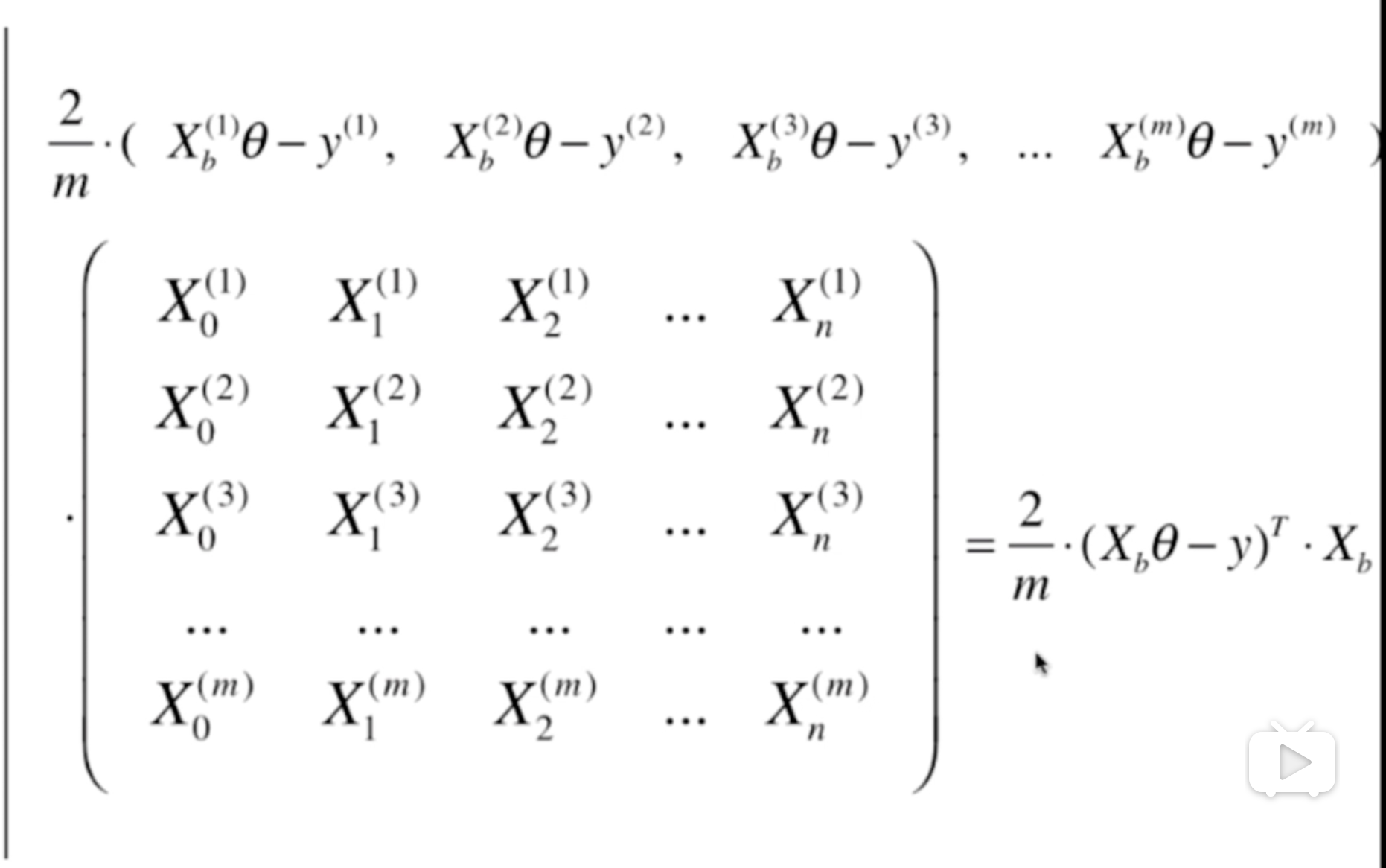
先把 X0 加上，让 X 和 theta 的长度对齐



在上一小节中，在for循环中已经对每一个偏导数的计算一定程度的用到了向量化，即把每一个偏导数的计算看做了两个向量的点乘

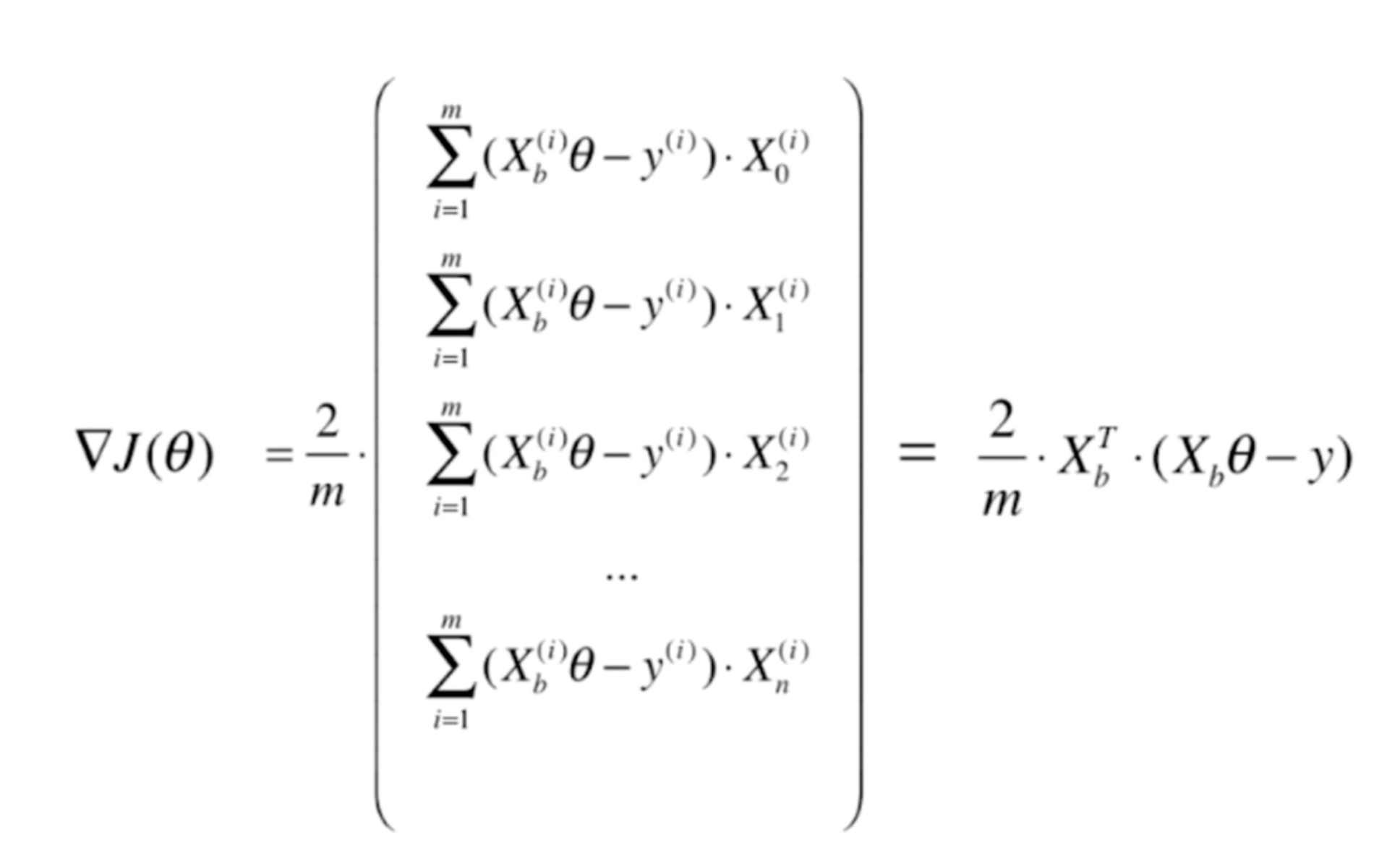
如果把外面的for循环去掉，那整个梯度的式子就可以进一步向量化为两个矩阵的点乘运算：

### 上节中我们看到：(Xbθ-y) 部分是一个 （m，1）的向量，为了参与后面的矩阵运算，把它装置，(Xbθ-y)T（上面截图的右上角一行）是一个 （1，m）的向量，下面部分就是一个 (m，n+1) 的矩阵，最终得到一个（1，n+1）的向量，正好满足待求的梯度中的 n+1 个元素



所以，计算的结果就可以用这么简单的矩阵。

可是我们上面得到的结果是一个 （1，n+1）的行向量，但是我们的梯度其实是一个（n+1，1）的列向量，才方便参与后面的计算。虽然 numpy中向量的运算不区分行向量和列向量，但是严谨起见，还是转换为列向量。所以最外面还要加一次装置，根据公司 (A\*B)T = BT \* AT，就得到公式：



把之前的代码中的 for 循环替换成行的向量化代码：[06-05 Vectorize-Gradient-Descent.ipynb]