Foundations of Data Science & Machine Learning

Tutotial 09

May 6, 2021

Question 1. Manually determine the best fit line through the origin for each of the following sets of points. Is the best fit line unique? Justify your answers in each case.

1. $\{(0,1),(1,0)\}$
2. $\{(0,1),(2,0)\}$

Solution 1.1 Let θ be the angle subtended by the best-fit line (through origin) S with the x-axis. So $d((1,0),S)$, the perpendicular distance from $(1,0)$ to the line S, is $\sin\theta$. Similarly $d((0,1),S) = \sin(90-\theta) = \cos\theta$. Please draw a picture and use high-school trigonometry to verify the above claims. The cost function we are minimising is $d^2((0,1),S) + d^2((1,0),S) = sin^2\theta + \cos^2\theta$ which is 1 for any value of θ. Hence every line through origin is equally the best.

You can also see the same by noticing that if you order the two points correctly then the data matrix A becomes the 2×2 identity matrix. Hence every unit vector v gives the same value for $\|Av\|$.

Solution 1.2 Let θ be the angle subtended by the best-fit line (through origin) S with the x-axis. So $d((2,0),S) = 2\sin\theta$. $d((0,1),S) = \sin(90-\theta) = \cos\theta$. Please draw a picture and use high-school trigonometry to verify the above claims. The cost function we are minimising is $d^2((0,1),S) + d^2((1,0),S) = 2sin^2\theta + \cos^2\theta = 1 + \sin^2\theta$ which is minimised when $\theta = 0 \pmod{\pi}$. Hence the best-fit line is the x-axis and it is unique.

Question 2. Let

$$C_{p\times r} = A_{p\times q}B_{q\times r}.$$

Prove that

$$C = \sum_{i=1}^{q} A[:,i]B[i,:],$$

where $A[:,i]$ denotes the i-th column of A and $B[i,:]$ denotes the i-th row of B. (Hence $A[:,i]B[i,:]$ is an outer product.)

Solution The (s,t)-the entry in $\sum_{i=1}^{q} A[:,i]B[i,:]$ is $\sum_{i=1}^{q} A[s,i]B[i,t]$ which is the same as the standard matrix multiplication formula.

Question 3. Let A be a square $n\times n$ matrix whose rows are orthonormal. Prove that the columns of A are orthonormal. *Hint.* Projection operator.

Solution. It is easy to verify by the inner-product view of matrix multiplication that $AA^T = I$. Let $a_1 \dots, a_n$ denote the rows of A which are given to be orthonormal. Hence $a_1,\dots,a_n$ form an orthonormal basis for $\mathbb{R}^n$. Hence A^TA is the projection operator for the subspace $\mathbb{R}^n$ itself. Since the projection operator on the entire vector space is the identity operator, $A^TA = I$. This shows that the columns of A are also orthonormal.