Algorithmic Foundations of Data Science: Assignment #8

Khaled AlHosani (kah579), Myunggun Seo (ms9144)

New York University Abu Dhabi
April 15, 2018

Problem 1

Given: A and B are two matrices s.t. $A\boldsymbol{v} = B\boldsymbol{v}$ for any vector $\boldsymbol{v}$.
Set up vectors $\boldsymbol{v}_i = (\mathbb{1}_i)$ which are vectors of dimensionality equal to number of columns of A and B and have 0 everywhere except for the i-th component, which is 1. If $A\boldsymbol{v_i} = B\boldsymbol{v_i}$ for every i from 1 to the number of columns of A and B, we have that $A\boldsymbol{v_i}$, which is a column in A, is equal to $B\boldsymbol{v_i}$, which is a column in B at the same location. Since every column of A and B are equal, A and B must be equal.

Problem 2

We prove by construction that it is possible to choose an orthonormal basis $\boldsymbol{w}_1, ..., \boldsymbol{w}_{k+1}$ for W such that $\boldsymbol{w}_{k+1} \perp \boldsymbol{v}_1, ..., \boldsymbol{v}_k$: for V of dimension k and W of dimension $k+1$, the projection of V onto W is a subspace between 0 and k dimensions. Since this projection is a subspace of lower dimension than W, we can always find a vector $\boldsymbol{w}_{k+1}$ in W that is orthogonal to this projection. This vector is orthogonal to V because it is orthogonal both to the projection of V and perpendicular of V on W. (W, which includes $\boldsymbol{w}_{k+1}$, is orthogonal to the perpendicular of V on W.)

By definition $|A\boldsymbol{w}_1| + ... + |A\boldsymbol{w}_{k+1}|$ is maximized for orthonormal vectors since W is the best fit $(k+1)$-dimensional subspace. Additionally, $|A\boldsymbol{v}_1| + ... + |A\boldsymbol{v}_k|$ is similarly maximized since V is the best fit k-dimensional subspace.

Which is to say that:

$|A\boldsymbol{w}_1| + ... + |A\boldsymbol{w}_k| \leq |A\boldsymbol{v}_1| + ... + |A\boldsymbol{v}_k|$

If we replace $\boldsymbol{w}_1 ... \boldsymbol{w}_k$ by $\boldsymbol{v}_1 ... \boldsymbol{v}_k$ in W we get:

$|A\boldsymbol{w}_1| + ... + |A\boldsymbol{w}_{k+1}| \leq |A\boldsymbol{v}_1| + ... + |A\boldsymbol{v}_k| + |A\boldsymbol{w}_{k+1}|$

This shows that the subspace V' spanned by $\boldsymbol{v}_1, ..., \boldsymbol{v}_k$ and w_{k+1} is at least as good as W.

Problem 3

Let h be the k-dimensional best-fit subspace of n points in $\mathbb{R}^d$.
Let S be a subspace of dimension $d - k$ orthogonal to h.

Let p^* be the orthogonal projection of h onto S.
For any point p in $\mathbb{R}^d$, let p^* be the projection of p on S.
The distance between p^* and h^* is equal to the distance between p and h.

The best-fit surface h is one that minimizes $\sum d(p,h)^2$. Each $d(p,h)$ is equal to $d(p^*,h^*)$ so we have that, by substitution, h produces the minimum value of $\sum d(p^*,h^*)^2$.

In 1-dimension, the only value of h^* that gives the minimum value of $\sum d(p^*,h^*)^2$ is the centroid of all points p^*. Thus h^* is the centroid of p^*'s.

Now let c be the centroid of all points and c^* its projection on S. the projection c^* of centroid is equal to the centroid h^* of projections because a centroid is computed component-wise (we can use the bases of h and S). Thus h^* is the projection of the centroid.

Since h is the $d-k$ dimension subspace that is orthogonal to S, all h is projected onto $h*$ but no other subspace is. Thus the centroid c, of which the projection is equal to h^*, must be in h as well.