EPFL
School of Computer and Communication Sciences
Nicolas Flammarion & Martin Jaggi
www.epfl.ch/labs/mlo/machine-learning-cs-433

Problem Set 12, Dec 3, 2020 (Solutions to SVD Theory Questions)

Problem 1 (How to compute U and S efficiently):

In class, we saw that solving the eigenvector/value problem for the matrix $\boldsymbol{X}\boldsymbol{X}^\top$ gives us a way to compute $\boldsymbol{U}$ and $\boldsymbol{S}$. But in some instances $D \gg N$. In those cases, is there a way to accomplish this computation more efficiently?

Solution 1:

Consider the $N \times N$ matrix $\boldsymbol{X}^\top\boldsymbol{X}$. Similarly as before, we have

$$\boldsymbol{X}^\top\boldsymbol{X} = \boldsymbol{V}\boldsymbol{S}^\top\boldsymbol{S}\boldsymbol{V}^\top.$$

Let $\boldsymbol{v}_i$, $i = 1, \cdots, D$, denote the columns of $\boldsymbol{V}$. Then

$$\boldsymbol{X}^\top\boldsymbol{X}\boldsymbol{v}_j = \boldsymbol{V}\boldsymbol{S}^\top\boldsymbol{S}\boldsymbol{V}^\top\boldsymbol{v}_j = s_j^2\boldsymbol{v}_j. \tag{1}$$

So we see that the j-th column of $\boldsymbol{V}$ is an eigenvector of $\boldsymbol{X}^\top\boldsymbol{X}$ with eigenvalue s_j^2. Therefore, solving the eigenvector/value problem for the matrix $\boldsymbol{X}^\top\boldsymbol{X}$ gives us a way to compute $\boldsymbol{V}$ and $\boldsymbol{S}$.

Now multiply the identity (1) from the left by the matrix $\boldsymbol{X}$. We get

$$\boldsymbol{X}\boldsymbol{X}^\top(\boldsymbol{X}\boldsymbol{v}_j) = s_j^2(\boldsymbol{X}\boldsymbol{v}_j).$$

We see therefore that $\boldsymbol{u}_j = \boldsymbol{X}\boldsymbol{v}_j$ and so we can compute the desired eigenvectors $\boldsymbol{u}_j$ from the eigenvectors $\boldsymbol{v}_j$ without having to solve the $D \times D$ eigenvector/value problem.

Problem 2 (Positive semi-definite):

Show that if $\boldsymbol{X}$ is a $N \times N$ symmetric matrix then the SVD has the form $\boldsymbol{U}\boldsymbol{S}\boldsymbol{U}^\top$, where $\boldsymbol{U}$ is a $N \times N$ unitary matrix and $\boldsymbol{S}$ is a $N \times N$ diagonal matrix with non-necessarily positive entries. Show that if $\boldsymbol{X}$ is positive semi-definite, then all entries of $\boldsymbol{S}$ are non-negative.

Solution 2:

Consider $\boldsymbol{A} = \boldsymbol{X}\boldsymbol{X}^\top$ and $B = \boldsymbol{X}^\top\boldsymbol{X}$. By the SVD we know that $\boldsymbol{X} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^\top$. As we discussed in the course, the columns of $\boldsymbol{U}$ are eigenvectors of the first matrix and the columns of $\boldsymbol{V}$ are eigenvectors of the second matrix. But note that $\boldsymbol{A} = \boldsymbol{B}$ since $\boldsymbol{X}$ is symmetric. Hence the eigenspace associated to each distinct eigenvalue of $\boldsymbol{A}$ is equal to the eigenspace associated to the same eigenvalue of $\boldsymbol{B}$.

Set $\boldsymbol{U} = \boldsymbol{V}$ and let the columns of $\boldsymbol{U}$ be eigenvectors of $\boldsymbol{A}$. Compute $\boldsymbol{U}^\top\boldsymbol{X}\boldsymbol{V}$. This gives us a diagonal matrix which we can define to be $\boldsymbol{S}$. It's entries are not necessarily non-negative.

If the matrix is in addition positive semi-definite then the diagonal entries of $\boldsymbol{S}$ must in fact must be non-negative – multiplying the matrix from the left by $\boldsymbol{u}_j^\top$ and from the right by $\boldsymbol{u}_j$ gives s_j which must be non-negative if the quadratic form given by the matrix is assumed to be positive-definite.