

Problem Set 11, Dec 01, 2016 (Theory Questions, Tensorflow Introduction)

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

Consider the $N \times N$ matrix $X^\top X$. Similar to before we have

$$X^\top X = V S^\top S V^\top.$$

Let $v_i, i = 1, \dots, D$, denote the columns of V . Then

$$X^\top X v_j = V S^\top S V^\top v_j = s_j^2 v_j. \quad (1)$$

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

Now multiply the identity (1) from the left by the matrix X . We get

$$X X^\top (X v_j) = s_j^2 (X v_j).$$

We see therefore that $u_j = X v_j$ and so we can compute the desired eigenvectors u_j from the eigenvectors v_j without having to solve the $D \times D$ eigenvector/value problem.

Problem 2 (Positive semi-definite):

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

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

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