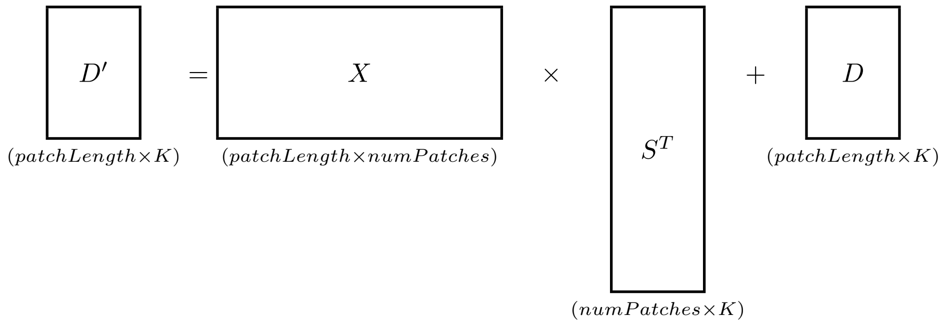
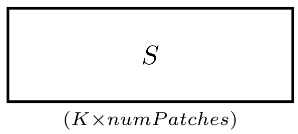
The matrices in spherical K-means

The update rule for the cluster centroids/dictionary D in spherical K-means can be shown diagrammatically like this:



The first variable in parentheses under each rectangle is the number of rows of the corresponding matrix, the second variable corresponds to the number of columns. K is the number of centroids (i.e., the number of dictionary elements/filters to be learned), and each column of D is a centroid. X is the whitened data matrix, with one patch/instance per column. The variable *patchLength* holds the number of pixels per patch, which is also the number of pixels per centroid, and the variable *numPatches* holds the total number of training patches/instances (including patches from all images).

The matrix S, whose transpose is S^T, has the format



Each column $s^{(i)}$ of this matrix S has exactly one non-zero element $s_j^{(i)}$, namely for the row j corresponding to the centroid $D^{(j)}$ that is most closely related to the image patch stored in column $x^{(i)}$ of X (as measured by the absolute value of the dot product between $D^{(j)}$ and $x^{(i)}$). The exact value of this non-zero element indicates how strongly the patch is aligned with the centroid. This value is the outcome of the dot product.

Note that you can calculate S very efficiently by using matrix multiplication instead of taking lots of individual dot products between column vectors in D and X. First, calculate S as

$$\underset{(K\times numPatches)}{\boxed{\quad S \quad}} = \underset{(K\times patchLength)}{\boxed{\; D^T \;}} \times \underset{(patchLength\times numPatches)}{\boxed{\quad X \quad}}$$

and then, in each column of S, keep only the element with the largest absolute value and set all other elements to 0.

PS: In the algorithm for spherical K-means, $\|p\|_2$ stands for the L_2 norm (aka Euclidean norm) of the vector p.