1 Additional Proof regarding Exercise 3.26 (a)

Proof:
\[
\sum_{i=1}^{k} \lambda_i(X) = \sup \{ tr(V^T X V) | V \in R^{n \times k}, V^T V = I \}
\]
where \( X \in S^n \) and
\[
\lambda_1(X) \geq \lambda_2(X) \geq \ldots \geq \lambda_n(X)
\]

Why is the variational characterization given in Boyd, Exercise 3.26 (a) correct?

**Proof.** First we show that such a solution for the supremum exists and then that it is indeed the maximum.

From the decomposition for symmetric matrices introduced in the first tutorial (and available in full detail on [www.socher.org](http://www.socher.org)), we know that any symmetric matrix \( X \) can be rewritten as:
\[
\Lambda = V^T X V 
\]
where \( \Lambda \) is the diagonal matrix with the eigenvalues on the diagonal. Hence, if we only align the \( k \) eigenvectors in \( V \) corresponding to the \( k \) largest eigenvalues, we get exactly a \( k \times k \) diagonal matrix with the \( k \) largest eigenvalues. Taking the trace of this matrix yields the given sum on the left side of equation 1.

Now the question arises, why the sum of the \( k \) largest eigenvalues corresponds to the supremum given in the right side of equation 1. Formally, is this correct:
\[
\{ tr(V^T X V) | V \in R^{n \times k}, V^T V = I \} \leq \sum_{i=1}^{k} \lambda_i(X)
\]

Let us take an arbitrary, rectangle, orthogonal matrix \( W \in R^{n \times k} \). We can represent this matrix \( W \) as columns of \( V \), since \( V \) spans the entire space in \( R^n \). Hence, there exists a \( Y \) such that
\[
W = V Y 
\]
so that
\[
tr(W^T X W) = tr(Y^T V^T X V Y) 
\]
\[
= tr(Y^T \Lambda Y) 
\]
\[
= \sum_{i=1}^{n} \lambda_i y_i^T y_i 
\]
Now, we can assume that \( Y \) is of rank \( k \) and its vectors are orthonormal and therefore
\[
\sum_{i=1}^{k} y_i^T y_i = k
\]
Another way to see this is: \( Y^T Y = Y^T V^T V Y = W^T W = I \in R^{k \times k} \) Using equation 2, we see:
\[
\sum_{i=1}^{n} \lambda_i y_i^T y_i \leq \sum_{i=1}^{k} \lambda_i
\]
\[
\square
\]