Abstract

We give algorithms for approximation by low-rank positive semidefinite (PSD) matrices. For symmetric input matrix $A \in \mathbb{R}^{n \times n}$, target rank k, and error parameter $\epsilon > 0$, one algorithm finds with constant probability a PSD matrix Y of rank k such that $\|A - Y\|_F^2 \leq (1 + \epsilon)\|A - A_{k,+}\|_F^2$, where $A_{k,+}$ denotes the best rank-k PSD approximation to A, and the norm is Frobenius. The algorithm takes time $O(\mathrm{nnz}(A) \log n) + n\mathrm{poly}((\log n)k/\epsilon) + \mathrm{poly}(k/\epsilon)$, where $\mathrm{nnz}(A)$ denotes the number of nonzero entries of A, and $\mathrm{poly}(k/\epsilon)$ denotes a polynomial in k/ϵ. (There are two different polynomials in the time bound.) Here the output matrix Y has the form $CUC^\top$, where the $O(k/\epsilon)$ columns of C are columns of A. In contrast to prior work, we do not require the input matrix A to be PSD, our output is rank k (not larger), and our running time is $O(\mathrm{nnz}(A) \log n)$ provided this is larger than $n\mathrm{poly}((\log n)k/\epsilon)$. We give a similar algorithm that is faster and simpler, but whose rank-k PSD output does not involve columns of A, and does not require A to be symmetric. We give similar algorithms for best rank-k approximation subject to the constraint of symmetry. We also show that there are asymmetric input matrices that cannot have good symmetric column-selected approximations.