Optimization problems in statistics[1]

T. Rapcsák[2]

Numerous statistical problems containing the principal component analysis, statistical visualization and singular value decomposition, furthermore, one of the basic theorems of linear algebra, the matrix spectral theorem and the characterization of the structural stability of dynamical systems are equivalent to the problem of finding, for any $2 \leq k \leq n$, the dominant k-dimensional eigenspace of an $n \times n$ symmetric matrix A in R^n where the eigenspaces are spanned by the corresponding eigenvectors. These problems lead to smooth nonlinear optimization problems with the same structural properties, namely, the feasible sets are Stiefel manifolds. This class, up to now, has insufficiently examined - in spite of their importance - either from theoretical or methodological point of view.

In 1935, Stiefel introduced a differentiable manifold [7], called Stiefel manifold, consisting of all the orthonormal vector systems $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k \in R^n$, where R^n is the n-dimensional Euclidean space and $k \leq n$. We consider the following optimization problem:

$$\begin{gathered}\min\ f(\mathbf{x}_1, \mathbf{x}_2, \ldots, \ldots, \mathbf{x}_k)\\ \mathbf{x}_i^T \mathbf{x}_j = \delta_{ij}, \qquad 1 \leq i, j \leq k \leq n,\\ \mathbf{x}_i \in R^n, \quad i = 1, \ldots, k, \qquad n \geq 2,\end{gathered}$$

where $f : R^{kn} \to R$ is a twice continuously differentiable function and δ_{ij} is the Kronecker's delta. Since the feasible set, the Stiefel manifold denoted by $M_{n,k}$ for every pair of positive integers (k, n), is compact and the objective function is continuous, the optimization problem has at least one global minimum point and one global maximum point, thus several stationary points.

This new and interesting global optimization problem with important theoretical and practical applications was studied in special cases in [1], [2], [3], [5]. In [6], the necessary and sufficient optimality conditions were obtained in the general case by the global Lagrange multiplier rule developed in [4].

In the lecture, the global optimality of the smooth nonlinear optimization problems originated in statistics are characterized based on the paper [6]. Since the background of the global Lagrange multiplier rule of smooth nonlinear optimization is the tensor approach in the corresponding Riemannian geometry, the proofs of the theorems in statistics are simple consequences of the dominant eigenspace theorem where no special coordinate representations are necessary.

[1] This research was supported in part by the Hungarian National Research Foundation, Grant No. OTKA-T029572
[2] Scientific Adviser, Computer and Automation Research Institute, Hungarian Academy of Sciences, H-1518 Budapest, P.O.Box 63,Hungary. E-mail: rapcsak@oplab.sztaki.hu
Internet: http://www.oplab.sztaki.hu/rapcsak/

References

(1) Bolla, M., Michaletzky, Gy., Tusnády, G. and Ziermann, M., Extrema of sums of heterogeneous quadratic forms, Linear Algebra and Applications 269 (1998) 331-365.

(2) Helmke, U. and Moore, J.B., Optimization and dynamical systems, Springer-Verlag, 1994.

(3) Márkus, L., Berke, O., Kovács, J. and Urfer, W., Spatial prediction of the intensity of latent effects governing hydrogeological phenomena, Environmetrics 10 (1999) 633-654.

(4) Rapcsák, T., Smooth nonlinear optimization in R^n, Kluwer Academic Publishers, Boston, London, Dordrecht, 1997.

(5) Rapcsák, T., On minimization of sums of heterogeneous quadratic functions on Stiefel manifolds, In: From local to global optimization, eds.: A. Migdalas, P. Pardalos and P. Varbrand, Dordrecht-Boston-London, Kluwer Academic Publishers. (in print)

(6) Rapcsák, T., On minimization on Stiefel manifolds, WP 2000-6, LORDS. (submitted for publication in EJOR)

(7) Stiefel, E., Richtungsfelder und Fernparallelismus in n-dimensionalen Mannigfaltigkeiten, Commentarii Mathematici Helvetici 8 (1935/6) 305-353.