

SVD Augmented Gradient Optimization

M.J. PAZDANOWSKI*

Cracow University of Technology, Faculty of Civil Engineering, Kraków, Poland

Solution time of nonlinear constrained optimization problem depends on the number of constraints, decision variables and conditioning of decision variables space. While the numbers of constraints and decision variables are external to the optimization procedure itself, one may try to affect the conditioning of the decision variables space within the self contained optimization module. This will directly affect the ratio of convergence of an iterative, gradient based optimization routine. Another opportunity for speedup of the solution process in case of quadratic objective function lies in the chance to eliminate the decision variables least affecting the objective function, and thus decrease the optimization problem size. Elimination of decision variables is based on the singular value decomposition of the objective function. Singular values showing up as a result of such procedure indicate that certain linear combinations of original decision variables do not affect the objective function, and thus may be eliminated from further deliberations. Also if near singular values are encountered as well, even deeper reduction of the optimization problem size is still possible, but at a cost in terms of final solution quality. An idea how to improve the conditioning of decision variables space, and limit the number of decision variables in case of quadratic objective function using singular value decomposition is presented in this paper. Results of computer tests performed during minimization of quadratic objective function and subject to quadratic constraints are enclosed and discussed.

DOI: [10.12693/APhysPolA.128.B-213](https://doi.org/10.12693/APhysPolA.128.B-213)

PACS: 02.60.Pn, 02.70.-c

1. Introduction

Solution time of an optimization problem strongly depends on the number of decision variables, constraints and conditioning of decision variables space. While the number of constraints may not be changed without affecting the final solution, one may try to affect the conditioning of decision variables space and the number of such variables via the application of proper numerical procedure. In the case of quadratic objective function (1) with positively defined matrix \mathbf{A} an application of $\mathbf{L} \cdot \mathbf{L}^T$ (Cholesky) decomposition accompanied by change of variables suffices [1], but this approach may not be used in the case of nonnegatively defined matrix \mathbf{A} . In such a situation an application of the singular value decomposition (SVD) [2] may help. In addition to improving the conditioning of the decision variables space this procedure may also let the user decrease the number of decision variables, shall the application of SVD indicate that matrix \mathbf{A} has one or more singular values. This in turn would mean that certain combinations of original decision variables do not affect the value of objective function, and as a result may be dropped from further analysis thus decreasing the size of decision variables space:

$$\min_{\mathbf{x}} F(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x}, \quad \mathbf{A} = \mathbf{A}^T. \quad (1)$$

One should note that though the SVD procedure itself [3–5] may be the single most time consuming element of the optimization process, it is performed only once at the beginning of calculations and the time spent on it is well offset by time gains on following calculations.

2. Solution approach

An application of SVD algorithm to square, symmetric matrix \mathbf{A} yields

$$\mathbf{A} = \mathbf{U}^T \mathbf{D} \mathbf{U}, \quad (2)$$

where \mathbf{U} is an orthonormal matrix ($\mathbf{U}^{-1} = \mathbf{U}^T$) and \mathbf{D} is a diagonal matrix containing moduli of eigenvalues of \mathbf{A} ordered by magnitude. If \mathbf{A} is rank deficient, at least one diagonal element of \mathbf{D} is equal to zero. Further decomposition of diagonal matrix \mathbf{D} into two components

$$\mathbf{D} = \mathbf{S} \mathbf{J} \mathbf{S}, \quad (3)$$

where \mathbf{S} is a diagonal matrix containing square roots of respective diagonal elements of \mathbf{A} with an exception of singular values, which are replaced by 1, and \mathbf{J} is a unit matrix, again with an exception of singular value locations in \mathbf{D} , which are replaced by 0. Thus the original objective function in the optimization problem (1) is replaced by

$$\min_{\mathbf{y}} F(\mathbf{y}) = \frac{1}{2} \mathbf{y}^T \mathbf{J} \mathbf{y}, \quad (4)$$

accompanied by an appropriate change of variables

$$\mathbf{y} = \mathbf{S} \mathbf{U} \mathbf{x}, \quad \mathbf{x} = \mathbf{S}^{-1} \mathbf{U}^T \mathbf{y}. \quad (5)$$

Based on the optimization problem at hand, one may either use formulae (5) to express the original constraints in new variables and switch all further calculations to the new, conditioned decision variables space, or work on two decision variable sets in parallel, should the complete transformation prove to be too cumbersome to perform. The second approach is tested in current paper, i.e. when the new vector of decision variables \mathbf{y} is determined, a corresponding vector \mathbf{x} is found using formulae (5), values of constraints and constraint gradients are computed in original decision variables space and transformed back to the conditioned space, where next iteration in search of the optimum solution is performed. One should note that the whole procedure is fully transparent to the end user of the optimization routines.

The algorithm of the above outlined procedure was implemented as a set of computer routines in Fortran programming language, supplemented by externally developed SVD routine [3].

*e-mail: plpazdan@cyfronet.krakow.pl

3. Test problem

A search for residual stress distribution in an elastic perfectly plastic body subjected to cyclic loads exceeding its elastic bearing capacity, based on [2], may be reduced to solving the nonlinear constrained optimization problem having quadratic objective function of the type (1) and quadratic constraints [7–10].

In order to check the accuracy and efficiency of the proposed extension, the test problem was solved several times with input data selected so, as to get optimization problems differing in size (number of decision variables and constraints), as well as in the decision variables to constraints ratio. Results representative for the tests performed so far are presented in Table.

Solution calculation times (in seconds).

TABLE

Case	Decision variables	Constraints	Iterations	Total time	Speed up	Time per iteration	Ratio	Global error [%]
STD	232	40	1753	39.20	–	0.0224	–	–
SVD1	232	40	327	28.30	1.39	0.0865	3.86	6.2515×10^{-4}
SVD2	137	40	310	21.60	1.81	0.0697	3.11	6.1927×10^{-4}
STD	256	47	5288	128.37	–	0.0243	–	–
SVD1	256	47	514	37.28	3.44	0.0725	2.98	2.9406×10^{-3}
SVD2	149	47	985	52.66	2.44	0.0535	2.20	2.9411×10^{-3}
STD	388	73	61262	1899.90	–	0.0310	–	–
SVD1	388	73	1754	368.48	5.16	0.2101	6.78	9.1649×10^{-3}
SVD2	220	73	1594	175.92	10.80	0.1104	3.56	9.1660×10^{-3}

In the table above STD denotes that the standard [11] set of optimization routines was used, while SVD1 denotes application of SVD augmented optimization procedure on full number of decision variables, and SVD2 denotes application of SVD augmented optimization routines including problem size reduction due to rank deficiency of matrix \mathbf{A} . Total time denotes time spent in optimization routines including the time needed to perform SVD decomposition when applicable.

The problem was solved on a 2.53 GHz Intel CoreDuo CPU equipped PC class computer.

Presented results seem to indicate that time gain due to application of SVD tends to grow with increasing problem size, with no loss in accuracy of final results incurred. This is confirmed by the relative difference between the reference (standard) and SVD augmented solution integrated over the whole domain and expressed as global error in the table above.

4. Conclusions

An idea to improve the convergence speed of iterative gradient solution method through conditioning of the decision variables space and, if possible, limiting the number of decision variables through application of SVD has been proposed. Tests performed on practical engineering problem [10] have shown that total calculation time gains varied between 40% and 980% depending on problem size with a very encouraging tendency to increase with increasing problem size.

Current application of the proposed procedure is limited to quadratic functions, but the proposed idea is quite general and the SVD decomposition may be applied to

any objective function, which may be locally approximated by a quadratic one. Unfortunately, this may be achieved only at the expense of efficiency, as the SVD decomposition, the single most time consuming step of the proposed procedure, would have to be applied after every change of the local approximation. This does require additional testing.

References

- [1] J. Orkisz, M. Pazdanowski, in: *Proc. X PCCM, Świnoujście (Poland)*, 1991, p. 591.
- [2] G.W. Stewart, *SIAM Rev.* **35**, 551 (1993).
- [3] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes in Fortran*, Cambridge University Press, Cambridge 1992.
- [4] *LAPACK Users' Guide*, SIAM, 1999.
- [5] *GNU Scientific Library Reference Manual*, 3rd ed., Network Theory, 2009.
- [6] J.B. Martin, *Plasticity — Fundamentals and General Results*, MIT Press, 1975.
- [7] W. Cecot, J. Orkisz, in: *Proc. COMPLAS V, Barcelona (Spain)*, 1997, p. 1879.
- [8] J. Orkisz, A. Harris, *Theor. Appl. Fract. Mech.* **9**, 109 (1988).
- [9] J. Orkisz, O. Orringer, M. Hołowiński, M. Pazdanowski, W. Cecot, *Comput. Struct.* **35**, 397 (1990).
- [10] M. Pazdanowski, *Arch. Transport* **XXII**, 319 (2010).
- [11] J. Orkisz, M. Pazdanowski, in: *The Finite Element Method in the 1990's*, Eds. E. Onate, J. Periaux, A. Samuelsson, Springer Verlag, 1991, p. 621.