Quadratic Kernels for Matrix Subspaces with Applications to Nonlinear Analysis and Optimization

Eugene Tyrtyshnikov

Institute of Numerical Mathematics of Russian Academy of Sciences
Lomonosov Moscow State University
Moscow Institute of Physics and Technology

eugene.tyrtyshnikov@gmail.com

Joint work with Alexey Tretyakov and Alexey Ustimenko

German-Russian Workshop, INM RAS
Moscow, June 14, 2017

Solving nonlinear equations

$$F(x) = 0, \quad x = [x_1, \dots, x_n]^\top, \quad F = [f_1(x), \dots, f_n(x)]^\top \in \mathbb{R}^n.$$

Consider the Jacoby matrix

$$F'(x) = [a_{ij}(x)], \quad a_{ij} = \partial f_i / \partial x_j.$$

Let $F(x^*) = 0$. If $F'(x^*)$ is nonsingular, then there are no other solutions in a vicinity of x^* and the Newton method converges quadratically:

$$x^{l+1} = x^l - (F'(x^l))^{-1} F(x^l), \quad l = 0, \ 1, \ \dots .$$

Further we assume that the matrix $F'(x^*)$ is of incomplete rank.

p-factor methods

Higner order derivatives (up to p) are used.

$$F^{(s)} = [a^{(s)}_{i_0,\, i_1,\ldots,i_s}], \quad a^{(s)}_{i_0,\, i_1,\ldots,i_s} = \frac{\partial^s f_{i_0}}{\partial x_{i_1} \ldots \partial x_{i_s}}$$

$$1 \leqslant s \leqslant p, \quad 1 \leqslant i_0, i_1, \ldots, i_s \leqslant n$$

For each value of i_0
we have a symmetric subtensor in $F^{(s)}$.

Multiplication of tensors

Given tensors $A^{(1)}, \dots, A^{(s)}$ with the elements

$$A^{(1)}_{i^1_1, \dots, i^1_{d_1}}, \quad \dots, \quad A^{(s)}_{i^s_1, \dots, i^s_{d_s}}$$

assume that in the whole list of indices $i_1, \dots, i_k$ occur only once and $j_1, \dots, j_l$ occur twice or more times. By *product of tensors* $A^{(1)}, \dots, A^{(s)}$ we mean a tensor B with the elements

$$B_{i_1, \dots, i_k} = \sum_{j_1, \dots, j_l} A^{(1)}_{i^1_1, \dots, i^1_{d_1}} \dots A^{(s)}_{i^s_1, \dots, i^s_{d_s}}.$$

We may write

$$B = A^{(1)} \dots A^{(s)}$$

but the detailed specification of indices is still needed.

Tretyakov's method

Let $p = 2$ and let Π denote a projector on a subspace complementary to $\mathrm{im} F'(x^*)$.

$$x^{l+1} = x^l - \underbrace{(F'(x^l) + \Pi F''(x^l)h)^{-1}}\underbrace{(F(x^l) + \Pi F'(x^l)h)}, \quad l = 0, 1, \ldots .$$

$F''h$ denotes a tensor $[b_{i_0,i_1}]$ which is a product of $F'' = [a_{i_0,i_1,i_2}]$ and $h = [h_{i_2}]$:

$$b_{i_0,i_1} = \sum_{i_2} a_{i_0,i_1,i_2} h_{i_2}.$$

Vector h is the same on all iterations and must provide nonsingularity of $F'(x^*) + \Pi F''(x^*)h$.
The mapping $x \to F(x)$ is called 2-regular at x^* if for some projector Π and some vector h the matrix $F'(x^*) + \Pi F''(x^*)h$ is nosingular. The above method is proved to converge quadratically.

Quadratic mappings

$$f_1(x) = \frac{1}{2}(Q_1 x, x), \quad \dots, \quad f_n(x) = \frac{1}{2}(Q_n x, x),$$

where $Q_1, \dots, Q_n$ are real symmetric matrices of order n. Since $F(0) = 0_n$ and $F'(0) = 0_{n \times n}$, the 2-regularity at $x^* = 0$ reads as follows: there exists $h \in \mathbb{R}^n$ s.t. the matrix

$$A = (F''h)^\top = [Q_1 h, \dots, Q_n h]$$

is nonsingular. Hence, h cannot be orthogonal to the span of $Q_1 h$, ..., $Q_n h$, that is,

$$h \notin \ker(F) = \{x \in \mathbb{R} : (Q_1 x, x) = \ldots = (Q_n x, x) = 0\}.$$

Now assume that the kernel of F is trivial and investigate the consequences.

Main result

Given a set $\mathcal{A}$ of real matrices of order n, define the *quadratic kernel* of $\mathcal{A}$ as the set of all $x \in \mathbb{R}^n$ s.t. $x^\top Ax = 0$ for any $A \in \mathcal{A}$.

THEOREM.
If a set of real symmetric matrices of order n has a trivial quadratic kernel, then its linear span contains a nonsingular matrix.

Corollaries

Note that the main result hold true even for symmetric matrices over sufficiently large finite fields.

COROLLARY. Let real symmetric matrices $Q_1, \ldots, Q_m$ of order n be such that any vector $x \in \mathbb{R}^n$ s.t.

$$x^\top Q_1 x = \ldots = x^\top Q_m x = 0$$

is zero. Then some linear combination

$$\alpha_1 Q_1 + \ldots + \alpha_m Q_m$$

with real coefficients is a nosingular matrix.

COROLLARY. The set of nonsingular matrices in the linear span of matrices $Q_1, \ldots, Q_m$ is dense in this span.

Absolute singularity

THEOREM. Assume that x^* satisfies

$$f_1(x^*) = \ldots = f_n(x^*) = 0$$

in the absolute singular case, that is,

$$f_1'(x^*) = \ldots = f_n'(x^*) = 0.$$

Let the quadratic kernel of the set $f_1''(x^*), \ldots, f_n''(x^*)$ is trivial. Then for some real numbers $\alpha_1, \ldots, \alpha_n$ the point x^* is a localy unique solution of the system

$$(f'(x))^\top e_1 = \ldots = (f'(x))^\top e_n = 0,$$
$$f(x) = \alpha_1 f_1(x) + \ldots + \alpha_n f_n(x).$$

$f'(x)$ is the column-gradient of the functional $f(x)$, $e_1, \ldots, e_n$ are the columns of the identity matrix of order n, and the Jacoby matrix at x^* for the blue system is nonsingular.

General singularity transformation

Let the rank of Jacoby matrix at x^* is equal to r and its first r rows be linearly independent. Then via a nonsingular transformation one can arrive to an equivalent system

$$\tilde{f}_1(x) = \cdots = \tilde{f}_n(x) = 0,$$

in which the first r equations copy the original equations

$$\tilde{f}_1(x) = f_1(x), \quad \ldots, \quad \tilde{f}_r(x) = f_r(x),$$

and $\tilde{f}_i$ for $r+1 \leqslant i \leqslant n$ is obtained from $f_i(x)$ by subtraction of some (depending on i) linear combination of $f_1(x), \ldots, f_r(x)$ providing us with the equalities

$$\tilde{f}'_{r+1}(x^*) = \ldots = \tilde{f}'_n(x^*) = 0.$$

General singularity theorem

THEOREM. Assume that

$$f'_{r+1}(x^*) = \ldots = f'_n(x^*) = 0,$$

the gradients $f'_1(x^*), \ldots, f'_r(x^*)$ are linearly independent, and the quadratic kernel for $f''_{r+1}(x^*), \ldots, f''_n(x^*)$ is trivial. Then for some real numbers $\alpha_{r+1}, \ldots, \alpha_n$ and vectors $h_{r+1}, \ldots, h_n \in \mathbb{R}^n$ the point x^* is a locally unique solution of the system

$$f_1(x) = 0, \quad \ldots, \quad f_r(x) = 0,$$

$$(f'(x))^\top h_{r+1} = 0, \quad \ldots, \quad (f'(x))^\top h_n = 0,$$

$$f(x) = \alpha_{r+1} f_{r+1}(x) + \ldots + \alpha_n f_n(x).$$

The Jacoby matrix of the blue system at x^* is nonsingular.

Optimality condition

Consider the optimization problem

$$\min \phi(x)$$

under constraint

$$F(x) = 0.$$

The Lagrange equation

$$\phi'(x^*) + (F'(x^*))^\top \lambda^* = 0$$

assumes that the matrix $F'(x^*)$ is of full rank.

What if the rank is not full?

THEOREM.
Let $\ker^2 Q = 0$. Then for any vector $y \in \mathbb{R}^n$ there are vectors $h \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}^m$ s.t.

$$y = (Qh)^\top \lambda.$$

COROLLARY
Let $\phi : \mathbb{R}^n \to \mathbb{R}$ and $F : \mathbb{R}^n \to \mathbb{R}^m$ be sufficiently smooth mappings and $m < n$. Assume the absolute singularity at x^* in the case $p = 2$ and the 2-kernel for $F^{(2)}(x^*)$ is trivial. Then there exist $h \in \mathbb{R}^n$ and $\lambda^* \in \mathbb{R}^m$ s.t.

$$\phi'(x^*) + (F''(x^*)h)^\top \lambda^* = 0.$$

The red condition may replace the Lagrange equation.

Seeking for nonsingular matrices in subspaces

THEOREM.

If a matrix of maximal rank in a matrix subspace is symmetric, then its kernel belongs to the quadratic kernel of this subspace.