МАТЕМАТИКА И МАТЕМАТИЧЕСКО ОБРАЗОВАНИЕ, 2009
MATHEMATICS AND EDUCATION IN MATHEMATICS, 2009
Proceedings of the Thirty Eighth Spring Conference of
the Union of Bulgarian Mathematicians
Borovetz, April 1–5, 2009

LYAPUNOV MAJORANTS FOR PERTURBATION ANALYSIS OF MATRIX EQUATIONS*

Mihail Konstantinov, Petko Petkov

We describe some efficient perturbation techniques for algebraic matrix equations. Among them are improved first order perturbation bounds, the method of equivalent operators and the technique of Lyapunov majorants combined with application of fixed point principles.

Introduction and notation. The sensitivity of computational problems is a major factor determining the accuracy of computations in machine arithmetic. It may be revealed and taken into account by the methods of perturbation analysis [14, 6]. Below we consider the technique of Lyapunov majorants for perturbation analysis of algebraic matrix equations $F(A, X) = 0$ arising in science and engineering, where A is a matrix parameter and X is the solution.

We shall use the following notations: $\mathrm{i} := \sqrt{-1}$ – the imaginary unit; $\mathbb{R}^{m\times n}$ and $\mathbb{C}^{m\times n}$ – the spaces of $m \times n$ matrices over the field of real $\mathbb{R}$ and complex $\mathbb{C}$ numbers; $\mathbb{R}^n = \mathbb{R}^{n\times 1}$, I_n – the identity $n\times n$ matrix; $\overline{A}$, $A^\top$ and $A^{\mathrm{H}} = \overline{A}^\top$ – the complex conjugate, the transpose and the complex conjugate transpose of the matrix A, respectively; $\mathrm{vec}(A)$ – the column–wise vectorization of the matrix A; $A \otimes B$ – the Kronecker product of the matrices A and B; $\|\cdot\|$ – a vector or a matrix norm; $\|\cdot\|_{\mathrm{F}}$ and $\|\cdot\|_2$ – the Frobenius norm and the 2–norm of a matrix or a vector, respectively; $V_n \in \mathbb{R}^{n^2\times n^2}$ – the vec–permutation matrix such that $\mathrm{vec}(Z^\top) = V_n\mathrm{vec}(Z)$ for $Z \in \mathbb{C}^{n\times n}$. The relation $\delta \succ 0$ ($\delta \succeq 0$) means that the real vector δ has positive (non-negative) elements, while the notation ‘:=’ means ‘equal by definition’.

Improved first order perturbation bounds. Suppose that the data A in the matrix algebraic equation

$$F(A, X) = 0 \tag{1}$$

in X is an m–tuple of matrices $A = (A_1, A_2, \ldots, A_m)$. Let these matrices be perturbed as $A_i \to A_i + E_i$, and let $X + Y$ be a solution of the perturbed equation $F(X + Y, A + E) = 0$, where $E = (E_1, E_2, \ldots, E_m)$. Using the Fréchet derivatives or pseudo-derivatives of the function F it is usually possible to derive expressions of the form $y \approx z := \sum_{i=1}^m L_i e_i$, where $y := \mathrm{vec}(Y)$, $e_i := \mathrm{vec}(E_i)$, and L_i are easily computable matrices. We note

*2000 Mathematics Subject Classification: 15A24, 15A21.
Key words: Perturbation analysis, Lyapunov majorants, algebraic matrix equations.

that $\|Y\|_F = \|y\|_2$. Let $\|E_i\|_F \le \delta_i$, where $\delta := [\delta_1, \delta_2, \dots, \delta_m]^\top \in \mathbb{R}_+^m$ is a given non–negative vector. For most problems we have $\|z - y\| = O(\|\delta\|^2)$, $\delta \to 0$, and hence $\|z\|_2$ approximates $\|Y\|_F$ up to first order terms in $\|\delta\|$.

Now the problem is to find a bound from above for $\|z\|_2$ as a function of δ, where the matrix $L := [L_1, L_2, \dots, L_m]$ is considered as a parameter.

The first estimate is $\|z\|_2 \le \mathrm{est}_1(L;\delta) = \mathrm{est}_1(L_1, L_2, \dots, L_m; \delta_1, \delta_2, \dots, \delta_m)$, where $\mathrm{est}_1(L;\delta) := \sum_{i=1}^m K_i\delta_i$ and $K_i := \|L_i\|_2$. This is a condition number based estimate since K_i is the absolute condition number of the problem relative to A_i.

Another immediate estimate is $\|z\|_2 \le \mathrm{est}_2(L;\delta) := \|L\|_2 \, \|\delta\|_2$. A third easily computable estimate [11] is $\|z\|_2 \le \mathrm{est}_3(L;\delta) := \sqrt{\delta^\top S \delta}$, where the elements s_{ij} of the matrix $S = [s_{ij}] \in \mathbb{R}_+^{m\times m}$ are defined from $s_{ij} := \left\|L_i^{\mathrm{H}} L_j\right\|_2$.

It may be shown that $\mathrm{est}_3(L;\delta) \le \mathrm{est}_1(L;\delta)$. Hence we have the perturbation result.

Theorem 1. *The following improved estimate is valid*

$$\|z\|_2 \le \mathrm{est}(L;\delta) := \min\{\mathrm{est}_2(L;\delta), \mathrm{est}_3(L;\delta)\}. \tag{2}$$

An interesting case arises in complex Lyapunov and Riccati equations, say

$$A_1 + A_2 X + X A_2^{\mathrm{H}} = 0, \tag{3}$$

where $A_1, A_2 \in \mathbb{C}^{n\times n}$ and $X \in \mathbb{C}^{n\times n}$. Suppose that $\lambda_i(A_2) + \overline{\lambda_j}(A_2) \ne 0$, $i, j = 1, 2, \dots, n$, where $\lambda_i(A_2)$ are the eigenvalues of the matrix A_2 counted according to their algebraic multiplicities. Under this assumption the matrix $A_0 := I_n \otimes A_2 + \overline{A_2} \otimes I_n \in \mathbb{C}^{n^2\times n^2}$ of the linear operator $X \mapsto A_2 X + X A_2^{\mathrm{H}}$ is non–singular and equation (3) has a unique solution X. Moreover, if $A_1^{\mathrm{H}} = A_1$ then $X^{\mathrm{H}} = X$ as well.

Let the coefficients and the solution of equation (3) be perturbed as $A_i \to A_i + E_i$, $X \to X + Y$, where $\|E_i\|_F \le \delta_i$, $i = 1, 2$. Denote $e_i := \mathrm{vec}(E_i)$ and $y := \mathrm{vec}(Y)$. Setting $A_3 := A_2^{\mathrm{H}}$ we see that $e_3 = \overline{e}_2$. Thus the perturbations e_2 and e_3 are not independent and a special technique to find tight perturbation bounds must be applied [7, 6]. The perturbed version of equation (3) is

$$A_1 + E_1 + (A_2 + E_2)(X + Y) + (X + Y)(A_2 + E_2)^{\mathrm{H}} = 0. \tag{4}$$

It may be shown that the inequality $\delta_2 < \delta_2^0 := 0.5\,\|A_0^{-1}\|_2^{-1}$ is sufficient for equation (4) in Y to have a unique solution. This condition is also "almost necessary" in the sense that for $\delta_2 = \delta_2^0$ the perturbed equation may have no solution or may have a variety of solutions.

It follows from (4) that $y = z + O(\|\delta\|^2)$, $\delta \to 0$, where $z := L_1 e_1 + L_2 e_2 + L_3 \overline{e}_2$ and $L_1 := -A_0^{-1}$, $L_2 := L_1(X^\top \otimes I_n)$, $L_3 := L_1(I_n \otimes X)V_n$. Here the function $e_2 \mapsto L_2 e_2 + L_3 \overline{e}_2$ in the expression for z is neither linear nor differentiable (it is additive but not homogeneous). Such functions and their norms have been studied in [6].

We have the inequality $\|z\|_2 \le \mathrm{est}(L_1, L_2, L_3; \delta_1, \delta_2, \delta_2)$ but this estimate is not tight enough. Instead, we may derive an improved estimate as shown below. Let $L_i := L_{i0} + \mathrm{i}L_{i1}$, $e_k := e_{k0} + \mathrm{i}e_{k1}$ and $z := z_0 + \mathrm{i}z_1$, where L_{ij}, e_{kj} and z_j are real. Denote by

$$z^{\mathbb{R}} := \begin{bmatrix} z_0 \\ z_1 \end{bmatrix}, \quad e_k^{\mathbb{R}} := \begin{bmatrix} e_{k0} \\ e_{k1} \end{bmatrix} \in \mathbb{R}^{2n}$$

the real versions of the vectors z and e_k, see [6]. Then we have $z^{\mathbb{R}} = M_1 e_1^{\mathbb{R}} + M_2 e_2^{\mathbb{R}}$, where

$$M_1 := \begin{bmatrix} L_{10} & -L_{11} \\ L_{11} & L_{10} \end{bmatrix}, \; M_2 := \begin{bmatrix} L_{20}+L_{30} & L_{31}-L_{21} \\ L_{21}+L_{31} & L_{20}-L_{30} \end{bmatrix} \in \mathbb{R}^{2n\times 2n}. \tag{5}$$

Now since $\|z^{\mathbb{R}}\|_2 = \|z\|_2$ we have the improved estimate

$$\|z\|_2 \le \mathrm{est}(M_1, M_2; \delta_1, \delta_2) \tag{6}$$

of type (2) which is better than the previous one.

Equivalent operator equation. Consider again the algebraic matrix equation (1) in X, where the matrix coefficients A_i are subject to perturbations $A_i \to A_i + E_i$, and let $X+Y$ be the solution of the perturbed equation

$$F(A+E, X+Y) = 0, \; E := (E_1, E_2, \ldots, E_m). \tag{7}$$

Suppose that the norms of the perturbations satisfy $\|E_i\|_{\mathrm{F}} \le \delta_i$, where $\delta_i \ge 0$ are given quantities. Then the aim of norm–wise perturbation analysis is to estimate the norm $\|Y\|_{\mathrm{F}}$ of Y as a function of the perturbation vector $\delta := [\delta_1, \delta_2, \ldots, \delta_m]^\top \in \mathbb{R}_+^m$,

Under some differentiability conditions for F the perturbed equation (7) may be written as an equivalent operator equation $Y = \Pi(E, Y)$, where

$$\begin{aligned} \Pi(E,Y) &:= -F_X^{-1}(A,X)(F_A(A,X)(E) + \; G(A,X,E,Y)), \\ G(A,X,E,Y) &:= F(A+E, X+Y) - F(A,X) - F_A(A,X)(E) - F_X(A,X)(Y). \end{aligned}$$

In turn, the matrix equation $Y = \Pi(E,Y)$ may be transformed into the vector equation $y = P(e,y)$, where $y := \mathrm{vec}(Y)$, $e := \mathrm{vec}(E)$ and $P(e,y) := \mathrm{vec}(\Pi(\mathrm{vec}^{-1}(e), \mathrm{vec}^{-1}(y)))$.

Further on, it may be shown that the operator $P(e,\cdot)$ transforms into itself a small set $\mathcal{B}_\rho$ of radius $\rho = f(\delta)$ vanishing together with δ. Thus according to the Schauder fixed point principle, there is a small solution for Y with $\|Y\|_{\mathrm{F}} \le f(\delta)$. The last inequality is the desired non–local perturbation estimate. This mechanism is described in the next section.

Lyapunov majorants. The technique of Lyapunov majorants goes back to the monographs [12, 2]. Of course, the first to use this technique was A.M. Lyapunov [13], see also [4]. Further developments on this subject may be found in [6]. Lyapunov majorants used in this paper are functions $(\delta,\rho) \mapsto l(\delta,\rho)$ described in the next three definitions.

Definition 2. *A function $l : \mathbb{R}_+^m \times \mathbb{R}_+ \to \mathbb{R}$ is said to be of* class Lyap *if it is continuous and non–decreasing in all its arguments, convex and differentiable in ρ and satisfies the conditions $l(0,0) = 0$, and $l'_\rho(0,0) < 1$.*

Consider again the matrix operator equation $Y = \Pi(E,Y)$ for the perturbation Y together with its vector counterpart $y = P(e,y)$, where $A = (A_1, A_2, \ldots, A_m)$ and $E = (E_1, E_2, \ldots, E_m)$, and let $\delta := [\delta_1, \delta_2, \ldots, \delta_m]^\top$ be a given non–negative vector.

Definition 3. *The function $l : \mathbb{R}_+^m \times \mathbb{R}_+ \to \mathbb{R}_+$, defined by the relation*

$$\begin{aligned} l(\delta,\rho) &:= \max\{\|\Pi(E,Y)\|_{\mathrm{F}} : \|E_i\|_{\mathrm{F}} \le \delta_i, \; \|Y\|_{\mathrm{F}} \le \rho\} \\ &= \max\{\|P(e,y)\|_2 : \|e_i\|_2 \le \delta_i, \; \|y\|_2 \le \rho\} \end{aligned}$$

is said to be the exact Lyapunov majorant *for the operator* Π *in the Frobenius norm.*

It may be shown that the function l from Definition 3 is of class *Lyap*. Moreover, for operators Π associated with non–linear algebraic matrix equations this function is non–linear and strictly convex in ρ.

Usually it is impossible to construct explicitly the exact Lyapunov majorant l. Instead, we use an easily computable function h which majorizes l in the sense that $h(\delta, \rho) \geq l(\delta, \rho)$.

Definition 4. *A function $h : \mathbb{R}_+^m \times \mathbb{R}_+ \to \mathbb{R}_+$ of class Lyap such that $h(\delta, \rho) \geq l(\delta, \rho)$ is said to be a* Lyapunov majorant *for the operator Π in the Frobenius norm.*

The technique of Lyapunov majorants is based on the *majorant equation*

$$\rho = h(\delta, \rho) \tag{8}$$

for determining ρ as a function of δ. Denote by $\Delta \subset \mathbb{R}_+^m$ the (non-empty) set of all $\delta \in \mathbb{R}_+^m$ such that equation (8) has a real non–negative solution. Since h is of class *Lyap* the set Δ has a non–empty interior Δ^o and for $\delta \in \Delta$ equation (8) has one or two solutions $\rho_1(\delta) \leq \rho_2(\delta)$ depending continuously on δ. Denoting $f(\delta) := \rho_1(\delta)$ we see that the solution $f(\delta)$ is small in the sense that the function $f : \Delta \to \mathbb{R}_+$ is continuous and $f(0) = 0$.

When the function $h(\delta, \cdot) : \mathbb{R}_+ \to \mathbb{R}_+$ is strictly convex for $\delta \succ 0$ fixed, the boundary of the domain Δ consists of the non–negative coordinate semi–planes in $\mathbb{R}_+^m$ and the surface Σ is defined by the condition that equation (8) has multiple roots. The surface Σ has codimension 1 in the set $\mathbb{R}_+^m$ of the perturbation vectors δ. To obtain Σ we should eliminate ρ from the system of two equations $\rho = h(\delta, \rho)$ and $1 = d(\delta, \rho)$, where $d(\delta, \rho) := \partial h(\delta, \rho)/\partial \rho$.

When the function h is affine in ρ, i.e. $h(\delta, \rho) = a_0(\delta) + a_1(\delta)\rho$ then equation (8) has the unique solution $f(\delta) = a_0(\delta)/(1 - a_1(\delta))$ provided $a_1(\delta) < 1$. The last inequality defines the domain Δ. Here the set Δ is not closed.

Consider now the more interesting case when the function h is non–linear (and hence strictly convex) in ρ. We may summarize our observations for this case as follows.

Theorem 5. *Let the Lyapunov majorant h be non–linear in ρ. Then*

(i) *the domain $\Delta \subset \mathbb{R}_+^m$ has a non–empty interior Δ^o;*

(ii) *for $\delta \in \Delta$ the majorant equation* (8) *has a small solution $\rho_1 = f(\delta)$ such that the function $f : \Delta \to \mathbb{R}_+$ is continuous, non-decreasing in all its arguments and $f(0) = 0$;*

(iii) *for $\delta \in \Delta^o$ the majorant equation* (8) *has two positive solutions $\rho_1(\delta) < \rho_2(\delta)$;*

(iv) *for the points $\delta \in \Sigma$ on the boundary of Δ it is fulfilled $\rho_1(\delta) = \rho_2(\rho)$.*

Denote $\mathcal{B}_\rho := \{y : \|y\|_2 \leq \rho\}$. For $\delta \in \Delta$ and $\|y\|_2 \leq f(\delta)$ we have $\|\Pi(E, Y)\|_{\mathrm{F}} = \|P(e, y)\|_2 \leq f(\delta)$. Thus the operator $P(e, \cdot)$ transforms the set $\mathcal{B}_{f(\delta)}$ into itself. Hence, according to the Schauder fixed point principle, there is a solution $y \in \mathcal{B}_{f(\delta)}$ of the operator equation $y = P(e, y)$. As a corollary we have the following important perturbation result.

Theorem 6. *For $\delta \in \Delta$, the perturbed equation* (7) *has a solution satisfying the non–local non–linear perturbation estimate $\|Y\|_{\mathrm{F}} \leq f(\delta)$.*

Hence the inclusion $\delta \in \Delta$ guarantees that the perturbed equation (7) is solvable and that the estimate $\|Y\|_{\mathrm{F}} \leq f(\delta)$ is rigorous.

In practice the domain Δ is not constructed explicitly. Rather, the inclusion $\delta \in \Delta$ is verified directly by checking a few explicit inequalities.

We note finally that for systems of matrix equations the resulting Lyapunov majorant is a vector valued function.

Linear equations. For linear matrix equations the Lyapunov majorant is affine in ρ, namely $h(\delta, \rho) = a_0(\delta) + a_1(\delta)\rho$, where the functions a_0, a_1 are non–negative, continuous, non–decreasing in δ and satisfy $a_0(0) = a_1(0) = 0$. Actually, these functions are of type est, described in Theorem 1. The domain Δ here is defined by the inequality $a_1(\delta) < 1$ and is not closed. The perturbation estimate is

$$\|Y\|_{\mathrm{F}} \leq \frac{a_0(\delta)}{1 - a_1(\delta)}, \ a_1(\delta) < 1.$$

Consider again equation (3) and its perturbed version (4). The equivalent vector operator equation for the perturbation $y = \mathrm{vec}(Y)$ here is

$$y = P(e, y) := L_1 e_1 + L_2 e_2 + L_3 \overline{e}_2 + L_1 \mathrm{vec}\left(E_2 Y + Y E_2^{\mathrm{H}}\right).$$

Therefore the Lyapunov majorant h is defined by

$$\|P(e, y)\|_2 \leq h(\delta, \rho) := \mathrm{est}(M_1, M_2; \delta_1, \delta_2) + 2\|L_1\|_2 \delta_2 \rho, \ \|y\|_2 \leq \rho,$$

where the matrices M_1, M_2 are given in (5). For $\delta_2 < \delta_2^0 := 0.5\,\|L_1\|_2^{-1}$ the majorant equation $\rho = h(\delta, \rho)$ has a unique solution $\rho = f(\delta)$. This result may be formulated as follows.

Theorem 7. *Let $\delta_2 < \delta_2^0$. Then the perturbed version* (4) *of the Lyapunov equation* (3) *has an unique solution Y so that the perturbation estimate*

$$\|Y\|_{\mathrm{F}} = \|y\|_2 \leq f(\delta) := \frac{\mathrm{est}(M_1, M_2; \delta_1, \delta_2)}{1 - 2\|L_1\|_2 \delta_2}$$

holds true.

Quadratic equations. Perturbation analysis of algebraic matrix quadratic equations has been done by many authors, see for example [9, 10, 5, 15], the monograph [6] and the references therein. For quadratic matrix equations $Q + \sum_i A_i X B_i + \sum_k C_k X D_k X E_k = 0$ the Lyapunov majorant is quadratic, $h(\delta, \rho) = a_0(\delta) + a_1(\delta)\rho + a_2(\delta)\rho^2$, where $a_0(\delta)$, $a_1(\delta)$ are expressions of type $\mathrm{est}(L; \delta)$. Hence the majorant equation is $a_2(\delta)\rho^2 - (1 - a_1(\delta))\rho + a_0(\delta) = 0$ and the domain Δ is defined by $\Delta = \{\delta \in \mathbb{R}_+^m : a_1(\delta) + 2\sqrt{a_0(\delta) a_2(\delta)}\} \leq 1\}$ (note that here Δ is a closed subset of $\mathbb{R}_+^m$). Thus we have established the following result.

Theorem 8. *For $\delta \in \Delta$, the corresponding perturbed matrix quadratic equation has a solution $X + Y$ such that Y satisfies the non–local perturbation estimate*

$$\|Y\|_{\mathrm{F}} \leq f(\delta) := \frac{2a_0(\delta)}{1 - a_1(\delta) + \sqrt{(1 - a_1(\delta))^2 - 4a_0(\delta)a_2(\delta)}}.$$

Consider, for example, the matrix equation

$$F(A, X) := A_1 + A_2 X + X A_3 + X A_4 X = 0,$$

where $A_1 \in \mathbb{C}^{m \times n}$, $A_2 \in \mathbb{C}^{m \times m}$, $A_3 \in \mathbb{C}^{n \times n}$, $A_4 \in \mathbb{C}^{n \times m}$ are given matrix coefficients and $X \in \mathbb{C}^{m \times n}$ is the solution. Let the matrices A_i be perturbed to $A_i \to A_i + E_i$ and $X + Y$ be a solution of the perturbed equation $F(A + E, X + Y) = 0$.

After elementary calculations, the perturbed equation can be written as

$$\begin{aligned}(A_2 + XA_4)Y + Y(A_3 + A_4X) &= -E_1 - E_2X - XE_4 - XE_4X \\ &\quad - E_2Y - YE_3 - XE_4Y - YE_4X - Y(A_4 + E_4)Y.\end{aligned}$$

Suppose that the linear matrix operator $\mathcal{H} : \mathbb{C}^{m\times n} \to \mathbb{C}^{m\times n}$ (the Fréchet derivative $F_X(A, X)$), defined by $\mathcal{H}(Y) := (A_2 + XA_4)Y + Y(A_3 + A_4X)$, is invertible. This is equivalent to the assumption that the matrix $H := I_n \otimes (A_2 + XA_4) + (A_3 + A_4X)^\top \otimes I_m$ of $\mathcal{H}$ is non–singular. Then we may rewrite the vectorized perturbed equation as $y = P(e, y) := P_1(e) + P_2(e, y) + P_3(e, y)$, where $y := \mathrm{vec}(Y)$, $e := [e_1^\top, e_2^\top, e_3^\top, e_4^\top]^\top$, $e_k := \mathrm{vec}(E_k)$,

$$\begin{aligned} P_1(e) &:= L_1e_1 + L_2e_2 + L_3e_3 + L_4e_4, \\ P_2(e, y) &:= L_1\mathrm{vec}(E_2Y + YE_3) + L_2\mathrm{vec}(YE_4) + L_3\mathrm{vec}(E_4Y), \\ P_3(e, y) &:= L_1\mathrm{vec}(Y(A_4 + E_4)Y) \\ \text{and } L_1 := -H^{-1},\ L_2 &:= L_1(X^\top \otimes I_m),\ L_3 := L_1(I_n \otimes X),\ L_4 := L_1(X^\top \otimes X). \end{aligned}$$

Remark. For notational convenience, we use the same symbols L_i for matrices different from these already used for the case of linear matrix equations; this will again be done in the case of fractional-affine equations.

Suppose that $\|y\|_2 \le \rho$. Then it follows from the expressions for P_k that

$$\begin{aligned} \|P_1(e)\|_2 &\le a_0(\delta) := \mathrm{est}(L_1, L_2, L_3, L_4; \delta_1, \delta_2, \delta_3, \delta_4), \\ \|P_2(e, y)\|_2 &\le \rho a_1(\delta) := \rho\, \mathrm{est}(L_1, L_2, L_3; \delta_2 + \delta_3, \delta_4, \delta_4), \\ \|P_3(e, y)\|_2 &\le \rho^2 a_2(\delta) := \rho^2 \|L_1\|_2(\|A_4\|_2 + \delta_4). \end{aligned} \tag{9}$$

Thus we have the following result.

Theorem 9. *The perturbation estimate for Y in Theorem 8 is valid, where the expressions $a_k(\delta)$ are determined by (9).*

Higher degree equations. Matrix algebraic equations (1) involving r-th degree expressions ($r > 2$) in the solution X give rise to Lyapunov majorants

$$h_r(\delta, \rho) := \sum_{k=0}^{r} a_k(\delta)\rho^k,\ \ \delta := [\delta_1, \delta_2, \dots, \delta_m]^\top \in \mathbb{R}_+^m,$$

which are polynomials in $\rho \ge 0$ of degree r. Here a_k are continuous non–negative non–decreasing functions in δ of type est (see Theorem 1) or polynomials in δ with non–negative coefficients satisfying the conditions $a_0(0) = 0$, $a_1(0) < 1$ and $a_r(\delta) > 0$ for some $\delta \succ 0$. In what follows we suppose that δ is small enough in order to guarantee the inequality $a_1(\delta) < 1$ since for $a_1(\delta) \ge 1$ the majorant equation $\rho = h_r(\delta, \rho)$ has no positive solutions.

Under these conditions and for small δ the majorant equation in ρ has a small positive solution $f_r(\delta)$ such that the Frobenius norm $\|Y\|_\mathrm{F}$ of the perturbation Y in the solution of the perturbed matrix equation (7) is bounded by the quantity $f_r(\delta)$. The function f_r is continuous, non–negative, non–decreasing in all its arguments and satisfies $f_r(0) = 0$. Moreover, the bound $\|Y\|_\mathrm{F} \le f_r(\delta)$ is valid for $\delta \in \Delta_r \subset \mathbb{R}_+^m$, where Δ_r is the domain of all δ for which the majorant equation has non–negative roots.

The boundary $\partial\Delta_r$ of the domain Δ_r is defined by the pair of equations $\rho = h_r(\delta,\rho)$, $1 = d_r(\delta,\rho)$ and the inequality $\delta \succeq 0$, where $d_r(\delta,\rho) := \partial h_r(\delta,\rho)/\partial\rho$. Hence for $\delta \in \partial\Delta_r$ either the discriminant of the algebraic equation $a_0(\delta) - (1 - a_1(\delta))\rho + a_2(\rho)\rho^2 + \cdots + a_r(\delta)\rho^r = 0$ in ρ is zero or $\delta_k = 0$ for some $k = 1, 2, \ldots, m$.

The domain Δ_r has a non–empty interior Δ_r^o. In particular the inclusion $\delta \in \Delta_r^o$ implies $\delta \succ 0$.

In general, there is no convenient explicit expression for $f_r(\delta)$ when $r > 2$. Therefore the problem is to find a tight easily computable upper bound $\widehat{f}_r(\delta)$ for $f_r(\delta)$. For this purpose and for small δ the r–th degree Lyapunov majorant $h_r(\delta,\rho)$ is replaced by a second degree Lyapunov majorant $\widehat{h}_r(\delta,\rho) := a_0(\delta)+a_1(\delta)\rho+b_r(\delta)\rho^2$ such that $\widehat{h}_r(\delta,\rho) \geq h_r(\delta,\rho)$ for $\rho \in [0,\tau(\delta)]$. Here $\tau(\delta)$ is a certain quantity which is positive for $\delta \succ 0$ and satisfies the inequality $h_r(\delta,\tau(\delta)) \leq \tau(\delta)$. Note that the first two terms in the expressions $h_r(\delta,\rho)$ and $\widehat{h}_r(\delta,\rho)$ coincide which guarantees that the use of $\widehat{h}_r(\delta,\rho)$ instead of $h_r(\delta,\rho)$ will produce a tight perturbation bound.

Denote by $\widehat{f}_r(\delta)$ the small solution of the new majorant equation $\rho = \widehat{h}_r(\delta,\rho)$. Then we obtain the perturbation estimate

$$\|Y\|_{\mathrm{F}} \leq \widehat{f}_r(\delta) := \frac{2a_0(\delta)}{1 - a_1(\delta) + \sqrt{(1 - a_1(\delta))^2 - 4a_0(\delta)b_r(\delta)}} \tag{10}$$

provided $a_1(\delta)+2\sqrt{a_0(\delta)b_r(\delta)} \leq 1$ and $\widehat{f}_r(\delta) \leq \tau(\delta)$ (or, equivalently, $h_r(\delta,\tau(\delta)) \leq \tau(\delta)$).

We stress that $a_0(\delta) = \mathrm{O}(\|\delta\|)$, $\delta \to 0$. Then both quantities $f_r(\delta)$ and $\widehat{f}_r(\delta)$ have asymptotic expansions $a_0(\delta)/(1 - a_1(0)) + \mathrm{O}(\|\delta\|^2)$, $\delta \to 0$. Hence we have the following result.

Theorem 10. *The asymptotic relation $\widehat{f}_r(\delta) = f_r(\delta) + \mathrm{O}(\|\delta\|^2)$, $\delta \to 0$, takes place.*

To find $\widehat{h}_r(\delta,\rho)$ and $\tau(\delta)$ we proceed as follows. For any $\tau > 0$ and $\rho \leq \tau$ we have

$$h_r(\delta,\rho) \leq g_r(\delta,\tau,\rho) := a_0(\delta) + a_1(\delta)\rho + \beta_r(\delta,\tau)\rho^2,$$

where

$$\beta_r(\delta,\tau) := a_2(\delta) + \sum_{k=2}^{r-1} a_{k+1}(\delta)\tau^{k-1}. \tag{11}$$

Let $\tau(\delta)$ be a positive non–decreasing expression in δ and $\rho \leq \tau(\delta)$. Then we may find a bound from above for $\beta_r(\delta,\tau(\delta))$, e.g. $b_r(\delta) \geq \beta_r(\delta,\tau(\delta))$, and use it in the estimate (10). Choosing different expressions $\tau(\delta)$ we obtain different bounds $b_r(\delta)$ for $\beta_r(\delta,\tau(\delta))$, different Lyapunov majorants $\widehat{h}_r(\delta,\rho)$ and, as a result, different bounds $\widehat{f}_r(\delta)$ for $\|Y\|_{\mathrm{F}}$. An useful observation here is that if the majorant equation $\rho = h_r(\delta,\rho)$ has positive solutions, then the small solution $f_r(\delta)$ does not exceed the value of ρ, where $d_r(\delta,\rho)$ reaches 1.

Consider the equation $1 = d_r(\delta,\rho)$, i.e.

$$1 = \sum_{k=0}^{r-1} (k+1)a_{k+1}(\delta)\rho^k, \tag{12}$$

in ρ. We have $d_r(\delta,0) = a_1(\delta) < 1$ for δ sufficiently small and $d_r(\delta,\rho) > 1$ for e.g. $\rho > (ra_r(\delta))^{1/(1-r)}$, and any $\delta \succ 0$. Hence for small $\delta \succ 0$ there is a unique positive solution $\rho = \tau_r(\delta)$ of equation (12). We stress that the solution $\tau_r(\delta)$ may exist even when the majorant equation $\rho = h_r(\delta, \rho)$ has no positive solution. But if the majorant equation has positive solutions $\rho_1(\delta) \le \rho_2(\delta)$ then $\rho_1(\delta) \le \tau_r(\delta) \le \rho_2(\delta)$ and $h_r(\delta, \tau_r(\delta)) \le \tau_r(\delta)$ by necessity.

It is clear from the above considerations that we may replace the quantity $\tau(\delta)$ by $\tau_r(\delta)$ or by some larger quantity. Furthermore, the terms $a_{k+1}(\delta)\tau_r^{k-1}(\delta)$ in (11) are bounded from above by suitable expressions, thus obtaining a new quadratic Lyapunov majorant. Below we describe this technique in more detail.

The case $r = 3$. Here $\tau_3(\delta)$ can be computed directly from the quadratic equation $3a_3(\delta)\rho^2 + 2a_2(\delta)\rho - (1 - a_1(\delta)) = 0$ as

$$\tau_3(\delta) = \frac{1 - a_1(\delta)}{a_2(\delta) + \sqrt{a_2^2(\delta) + 3a_3(\delta)(1 - a_1(\delta))}}. \tag{13}$$

For $\rho \le \tau_3(\delta)$ we have $h_3(\delta, \rho) \le \widehat{h}_3(\delta, \rho) := a_0(\delta) + a_1(\delta)\rho + b_3(\delta)\rho^2$, where $b_3(\delta) := a_2(\delta) + a_3(\delta)\tau_3(\delta)$. Hence $\widehat{h}_3(\delta, \rho)$ is a new Lyapunov majorant, which is quadratic in ρ. As a result we get the next result.

Theorem 11. *Let $a_1(\delta) + 2\sqrt{a_0(\delta)b_3(\delta)} \le 1$. Then we have the perturbation estimate*

$$\|Y\|_{\mathrm{F}} \le \widehat{f}_3(\delta) := \frac{2a_0(\delta)}{1 - a_1(\delta) + \sqrt{(1 - a_1(\delta))^2 - 4a_0(\delta)b_3(\delta)}} \tag{14}$$

provided that $\widehat{f}_3(\delta) \le \tau_3(\delta)$.

It may be shown that the inequality $\widehat{f}_3(\delta) \le \tau_3(\delta)$ in Theorem 11 is equivalent to $h_3(\delta, \tau_3(\delta)) \le \tau_3(\delta)$. Any of these inequalities is easily checkable in view of the explicit expression (13).

The case $r > 3$. Here the technique used is more involved since $\tau_r(\delta)$ may not be computed explicitly. Instead, we work with certain easily computable quantities $\alpha_{k+1}(\delta) \ge a_{k+1}(\delta)\tau_r^{k-1}(\delta)$ in (11), see [6].

Consider again equation (12) for a given small $\delta \succ 0$ which guarantees that the equation has a (unique) root $\tau_r(\delta) > 0$. This in particular implies $a_1(\delta) < 1$. For $k = 2, 3, \ldots, r-1$ we have $(k+1)a_{k+1}(\delta)\tau_r^k(\delta) \le 1 - a_1(\delta)$ and $\tau_r(\delta) \le ((1 - a_1(\delta)))/((k+1)a_{k+1}(\delta)))^{1/k}$ whenever $a_{k+1}(\delta) > 0$. Hence

$$a_{k+1}(\delta)\tau_r^{k-1}(\delta) \le \alpha_{k+1}(\delta) := a_{k+1}^{1/k}(\delta)\left(\frac{1 - a_1(\delta)}{k+1}\right)^{1-1/k}$$

and

$$\beta_r(\delta, \tau_r(\delta)) \le b_r(\delta) := a_2(\delta) + \sum_{k=2}^{r-1} \alpha_{k+1}(\delta).$$

Thus we have obtained the following perturbation result.

Theorem 12. *The perturbation estimate*

$$\|Y\|_{\mathrm{F}} \le \widehat{f}_r(\delta) := \frac{2a_0(\delta)}{1 - a_1(\delta) + \sqrt{(1 - a_1(\delta))^2 - 4a_0(\delta)b_r(\delta)}} \tag{15}$$

is valid provided $a_1(\delta) + 2\sqrt{a_0(\delta)b_r(\delta)} \le 1$ and $\widehat{f}_r(\delta) \le \min\{\alpha_{k+1}(\delta) : k = 2, 3, \ldots, r-1\}$.

Fractional–affine equations. Fractional–affine matrix equations involve inversions of affine expressions in X. Typical example here is the discrete–time matrix Riccati equation $Q - X + A^{\mathrm{H}} X(I + MX)^{-1} A = 0$ arising in optimal control and filtering of discrete–time systems, where the matrices $Q = Q^{\mathrm{H}}$ and $M = M^{\mathrm{H}}$ are non–negative definite, the pair $[A, M)$ is controllable, and the pair $(Q, A]$ is detectable. The Lyapunov majorants $h(\rho, \delta)$ for such equations may be chosen as quadratic polynomials or fractional–affine expressions in ρ.

Consider the following Lyapunov majorant arising from a certain fractional–affine matrix equation

$$h(\delta, \rho) := b_0(\delta) + b_1(\delta)\rho + \frac{b_2(\delta) + b_3(\delta)\rho + b_4(\delta)\rho^2}{b_5(\delta) - b_6(\delta)\rho}, \tag{16}$$

where $\delta \in \mathbb{R}_+^m$ and $b_k(\delta) \geq 0$. Suppose that a) the functions $b_1, b_2, \ldots, b_6$ are continuous, b) the functions b_k are non–decreasing in δ for $k \neq 5$, c) the function b_5 is positive and non-increasing in δ, and d) the relations $b_0(0) = b_2(0) = 0$, $b_1(0) < 1$, $b_5(0) > 0$, $b_6(0) > 0$ and $d(0, 0) = b_1(0) + b_3(0)/b_5(0) < 1$ take place, where $d(\delta, \rho) := \partial h(\delta, \rho)/\partial \rho$.

Denote

$$\begin{aligned} c_0(\delta) &:= b_2(\delta) + b_0(\delta)b_5(\delta),\ c_1(\delta) := b_5(\delta)(1 - b_1(\delta)) + b_0(\delta)b_6(\delta) - b_3(\delta), \\ c_2(\delta) &:= b_4(\delta) + b_6(\delta)(1 - b_1(\delta)). \end{aligned} \tag{17}$$

Then we have $c_0(0) = 0$ and $c_1(0) = b_5(0)(1 - d(0, 0)) > 0$. Hence for small $\delta \succeq 0$ it is fulfilled that $c_1(\delta) > 0$ and $c_1^2(\delta) > 4c_0(\delta)c_2(\delta)$.

The majorant equation $\rho = h(\delta, \rho)$ may be written as $c_2(\delta)\rho^2 - c_1(\delta)\rho + c_0(\delta) = 0$. Therefore we come to the following result.

Theorem 13. **(i)** *The set* $\Delta := \left\{\delta \in \mathbb{R}_+^m : c_1(\delta) > 0,\ c_1^2(\delta) \geq 4c_0(\delta)c_2(\delta)\right\}$ *has non–empty interior.*

(ii) *A bound for the Frobenius norm of the perturbation in the solution of the corresponding fractional–affine matrix equation is given by*

$$f(\delta) := \frac{2c_0(\delta)}{c_1(\delta) + \sqrt{c_1^2(\delta) - 4c_0(\delta)c_2(\delta)}},\ \delta \in \Delta. \tag{18}$$

Consider for example the matrix equation

$$F(A, X) := A_1 + A_2 X + X A_3 + A_4 X^{-1} A_5 = 0, \tag{19}$$

where $A := (A_1, A_2, A_3, A_4, A_5)$ and $A_i, X \in \mathbb{C}^{n\times n}$. As before, let the matrix coefficients A_i be perturbed to $A_i + E_i$ and let $X + Y$ be a solution of the perturbed equation $F(A + E, X + Y) = 0$, where $E := (E_1, E_2, E_3, E_4, E_5)$.

Suppose that $\|Y\|_{\mathrm{F}} \leq \rho$ and $\rho < \sigma := \|X^{-1}\|_2^{-1}$. Then the matrix $Z := X + Y$ is invertible, and $Z^{-1} = X^{-1} - X^{-1}YZ^{-1} = X^{-1} - Z^{-1}YX^{-1} = X^{-1} - X^{-1}YX^{-1} + X^{-1}YZ^{-1}YX^{-1}$. Moreover, we have $\|Z^{-1}\|_2 \leq (\sigma - \rho)^{-1}$.

The perturbation analysis presented below is based on the identity $F(A+E, X+Y) = F(A, X) + \mathcal{L}(Y) + F_0(E) + F_1(E, Y) + F_2(Y)$, where $\mathcal{L}(Y) := A_2 Y + Y A_3 - A_4 X^{-1} Y X^{-1} A_5$ and

$$\begin{aligned} F_0(E) &:= E_1 + E_2 X + X E_3 + A_4 X^{-1} E_5 + E_4 X^{-1} A_5 + E_4 Z^{-1} E_5, \\ F_1(E, Y) &:= E_2 Y + Y E_3 - A_4 X^{-1} Y Z^{-1} E_5 - E_4 Z^{-1} Y X^{-1} A_5, \\ F_2(Y) &:= A_4 X^{-1} Y Z^{-1} Y X^{-1} A_5. \end{aligned}$$

Suppose that the linear matrix operator $\mathcal{L} := F_X(A, X)$ is invertible and denote

$B := \left(X^{-1}A_5\right)^\top \otimes \left(A_4X^{-1}\right)$. Then the matrix $L := I_n \otimes A_2 + A_3^\top \otimes I_n - B$ of $\mathcal{L}$ is also non–singular. Hence the perturbed equation may be written as $y = P(e,y) := P_0(e) + P_1(e,y) + P_2(y)$, where $y := \mathrm{vec}(Y)$, $e := \mathrm{vec}(E)$, $e_k := \mathrm{vec}(E_k)$ and

$$\begin{aligned} P_0(e) &:= L_1e_1 + L_2e_2 + L_3e_3 + L_4e_4 + L_5e_5 + L_1\mathrm{vec}(E_4Z^{-1}E_5), \\ P_1(e,y) &:= L_1\mathrm{vec}(E_2Y + YE_3) - L_4\mathrm{vec}(E_4Z^{-1}Y) - L_5\mathrm{vec}(YZ^{-1}E_5) \\ &\quad + L_1B\mathrm{vec}(YZ^{-1}Z),\ P_2(y) := A_4X^{-1}YZ^{-1}YX^{-1}A_5. \end{aligned}$$

Here the matrices L_k are defined by $L_1 := -L^{-1}$, $L_2 := L_1(I_n \otimes X)$, $L_3 := L_1(I_n \otimes X)$, $L_4 := L_1\left(\left(X^{-1}A_5\right)^\top \otimes I_n\right)$, $L_5 := L_1\left(I_n \otimes \left(A_4X^{-1}\right)\right)$.

For $\|y\|_2 \le \rho < \sigma$ and after standard calculations we get

$$\begin{aligned} \|P_0(e)\|_2 &\le \mathrm{est}(L_1, L_2, L_3, L_4; \delta_1, \delta_2, \delta_3, \delta_4) + \frac{\lambda\delta_4\delta_5}{\sigma - \rho}, \\ \|P_1(e,y)\| &\le \rho\lambda(\delta_2 + \delta_3) + \frac{\rho\,\mathrm{est}(L_4, L_5; \delta_4, \delta_5)}{\sigma - \rho},\ \|P_2(y)\|_2 \le \frac{\beta\rho^2}{\sigma - \rho}, \end{aligned}$$

where $\lambda := \|L_1\|_2$ and $\beta := \|L_1B\|_2$. These inequalities give a Lyapunov majorant of type (16) with $b_0(\delta) := \mathrm{est}(L_1, L_2, L_3, L_4; \delta_1, \delta_2, \delta_3, \delta_4)$, $b_1(\delta) := \lambda(\delta_2+\delta_3)$, $b_2(\delta) := \lambda\delta_4\delta_5$, $b_3(\delta) := \mathrm{est}(L_4, L_5; \delta_4, \delta_5)$, $b_4 := \beta$, $b_5 := \sigma$, $b_6 := 1$. Hence $c_0(0) = 0$, $c_1(0) = \sigma > 0$, $c_2(0) = 1 + \beta$, the condition $c_1^2(0) > 4c_0(0)c_2(0)$ is fulfilled and the domain Δ is correctly defined. Hence we may formulate our last result as follows.

Theorem 14. *The perturbation bound* (18), (17) *is valid for equation* (19) *with the values of $b_k(\delta)$ presented above.*

REFERENCES

[1] G. Golub, C. Van Loan. Matrix Computations. John Hopkins Univ. Press, Baltimore, 1996, ISBN 0-8018-5414-8, Zbl 865.65009.

[2] E. Grebenikov, Yu. Ryabov. Constructive Methods for Analysis of Nonlinear Systems. Nauka, Moscow, 1979 (in Russian).

[3] N. Higham, M. Konstantinov, V. Mehrmann, P. Petkov. The sensitivity of computational control problems. *IEEE Control Syst. Magazine*, **24** (2004), 28–43.

[4] L. Kantorovich. Principle of majorants and Newton's method. *Dokl. AN SSSR*, **76**, No 1 (1951), 17–20 (in Russian).

[5] C. Kenney, G. Hewer. The sensitivity of the algebraic and differential Riccati equations. *SIAM J. Control Optim.*, **28** (1990), 50–69, Zbl 695.65025.

[6] M. Konstantinov, D. Gu, V. Mehrmann, P. Petkov. Matrix Perturbation Theory. Elsevier, Amsterdam, 2003, ISBN 0-444-51315, MR 2004g:15015.

[7] M. Konstantinov, P. Petkov. Note on "Perturbation theory for algebraic Riccati equations" [*SIAM J. Matrix Anal. Appl.*, **19** (1998), 39–65, MR 99d:93016] by J.G. Sun. *SIAM J. Matrix Anal. Appl.*, **21** (1999), 327, MR 2000g:93034.

[8] M. Konstantinov, P. Petkov. The method of splitting operators and Lyapunov majorants in perturbation linear algebra and control. *Numer. Func. Anal. Appl.*, **23** (2002), 529–572, MR 2004d:93051.

[9] M. Konstantinov, P. Petkov, N. Christov. Perturbation analysis of the continuous and discrete matrix Riccati equations. Proc. 1986 Amer. Control Conf., Seattle, 1986, 636–639.

[10] M. Konstantinov, P. Petkov, N. Christov. Perturbation analysis of matrix quadratic equations. *SIAM J. Sci. Stat. Comput.*, **11** (1990), 1159–1163, MR 91m:15021.
[11] M. Konstantinov, P. Petkov, D. Gu. Improved perturbation bounds for general quadratic matrix equations. *Numer. Funct. Anal. Opim.*, **20** (1999), 717–736, MR 2000g:65035.
[12] D. Lika, Yu. Ryabov. Iterative Methods and Lyapunov Majorant Equations in Non-Linear Osillation Theory. Shtiinca, Kishinev, 1974 (in Russian).
[13] A. Lyapunov. The General Problem of Stability of Motion. Gostehizdat, Moscow, 1950 (in Russian).
[14] G. Stewart, J. Sun. Matrix Perturbation Theory. Academic Press, New York, 1990, ISBN 0-12-670230-6.
[15] J. Sun. Perturbation theory for algebraic Riccati equations. *SIAM J. Matrix Anal. Appl.*, **19** (1998), 39–65, Zbl 914.15009.

Mihail Konstantinov
UACEG
1046 Sofia, Bulgaria
e-mail: mmk_fte@uacg.bg

Petko Petkov
TU-Sofia
1756 Sofia, Bulgaria
e-mail: php@tu-sofia.bg

МАЖОРАНТИ НА ЛЯПУНОВ ЗА ПЕРТУРБАЦИОНЕН АНАЛИЗ НА МАТРИЧНИ УРАВНЕНИЯ

Михаил Константинов, Петко Петков

Описани са някои ефективни техники за пертурбационен анализ на матрични уравнения: подобрени пертурбационни граници от първи ред, метод на еквивалентните оператори и мажоранти на Ляпунов, в съчетание с прилагане на принципите на неподвижната точка.