ON A RATIONAL MATRIX EQUATION OCCURING IN STOCHASTIC CONTROL

T. Damm, D. Hinrichsen

Institute for Dynamical Systems
University of Bremen
28334 Bremen, Germany
Fax : +49 421 218 4235, e-mail : tobias@math.uni-bremen.de, dh@math.uni-bremen.de

Keywords: Riccati equations, resolvent positive operators, concavity, stochastic control, Newton's method.

Abstract

We regard a general class of rational matrix equations, which contains the continuous (CARE) and discrete (DARE) algebraic Riccati equations as special cases. Equations of this type were encountered in [EHP96] and [EHP98] where H^∞-type problems of disturbance attenuation for stochastic linear systems were studied. We develop a unifying framework for the analysis of these equations based on the theory of (resolvent) positive operators and show that they can be solved by Newton's method starting at an arbitrary stabilizing matrix.

1 Introduction

Important questions in optimal linear control theory can be formulated as minimization problems for quadratic functionals constrained by a linear system. The solution of these problems leads to certain matrix equations and inequalities (for a Hermitian matrix X) that differ depending both on the type of quadratic functionals regarded (semidefinite or indefinite) and on the type of system in question (continuous or discrete, deterministic or stochastic). If they exist, the largest solutions of these equations yield the optimal feedback gain matrix and thus the minimizing control.

Positive semidefinite functionals are regarded in classical LQ-control theory, whereas indefinite functionals occur e.g. in H^∞-optimal control. In the deterministic case they lead to *definite* and *indefinite* continuous and discrete algebraic Riccati equations (e.g. [LR95]), respectively. It turns out that the solvability question is much more delicate for the indefinite than for the definite matrix equations.

The incorporation of multiplicative noise in the above problems leads to Riccati type rational matrix equations; they comprise the previous ones as special cases but in general they contain additional terms, that are monotonous in X and stem from the diffusion part of the system; see [Won68] for the definite case and [EHP96], [EHP98] and also Section 2 for the indefinite case.

In the *deterministic* case there exist various approaches to the solution of the corresponding definite and indefinite matrix equations, and it is in many cases possible to convert a result for a continuous equation to a result for the discrete equation. Nevertheless both cases usually need to be treated individually.

In [Won68] an iterative Newton-type method has been applied to the *definite stochastic* equation. In the present paper we develop Newton's method to solve the *indefinite stochastic* equation. Actually we dispense with all definiteness assumptions and consider an equation that comprises all those equations mentioned above as special cases. In Section 5, we prove as our main result, that the Newton iteration starting from any stabilizing matrix converges to a solution of this equation if it is solvable at all.

In the proof of our main result we use methods and results from the theory of positive and resolvent positive operators which we present in Section 3 without proofs.

2 A bounded real lemma of stochastic control

Regard the linear Itô differential equation

$$\begin{aligned} dx(t) &= Ax(t)dt + Bv(t)dt \\ &\quad + \sum_{i=1}^{N} A_0^i x(t) dw_i(t) + \sum_{i=1}^{N} B_0^i(t) v(t) dw_i(t) \\ z(t) &= Cx(t) + Dv(t), \end{aligned} \tag{1}$$

where $(A, C) \in \mathbb{K}^{n\times n} \times \mathbb{K}^{q\times n}$, and

$$(A_0^i, B_0^i, B, D) \in \mathbb{K}^{n\times n} \times \mathbb{K}^{n\times \ell} \times \mathbb{K}^{n\times \ell} \times \mathbb{K}^{q\times \ell}.$$

The $(w_i(t))_{t\in\mathbb{R}_+}$ are independent zero mean real Wiener processes on a probability space $(\Omega, \mathcal{F}, \mu)$ with respect to an increasing family $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ of σ-algebras $\mathcal{F}_t \subset \mathcal{F}$.

Let $L^2_w(\mathbb{R}_+, \mathbb{K}^\ell)$ denote the corresponding space of non-anticipating stochastic processes v with

$$\|v(\cdot)\|^2_{L^2_w} := \mathcal{E}\left(\int_0^\infty \|v(t)\|^2 dt\right) < \infty,$$

where $\mathcal{E}$ denotes expectation.
It is known from Itô-theory, that for all $(x_0, v) \in \mathbb{K}^n \times L^2_w(\mathbb{R}_+, \mathbb{K}^\ell)$ there exists a unique solution $x(\cdot, x_0, v)$ of (1) with initial value $x(0, x_0, v) = x_0$ and thus also a unique output process $z(\cdot, x_0, v)$.
We write $z(\cdot, 0, v) = \boldsymbol{L}v(\cdot)$, and call $\boldsymbol{L}$ the *perturbation operator* of the system (1). It describes the effect of the input process v (viewed as a stochastic disturbance) on the output process z (interpreted as the vector of the to be controlled variables).

Definition 2.1 *The system (1) is said to be* internally (exponentially mean square) stable *if $x(\cdot, x_0, 0) \in L^2_w(\mathbb{R}_+, \mathbb{K}^\ell)$ for all $x_0 \in \mathbb{K}^n$, where $x(\cdot) = x(\cdot, x_0, 0)$ is the solution of the unperturbed system (with $v(\cdot) \equiv 0$).*
The system (1) is called externally stable *if $\boldsymbol{L}$ is a bounded operator $\boldsymbol{L}$: $L^2_w(\mathbb{R}_+, \mathbb{K}^\ell) \to L^2_w(\mathbb{R}_+, \mathbb{K}^q)$.*

In [EHP96] it was shown, that internal stability of (1) implies external stability. Then the norm $\|\boldsymbol{L}\|$ of the disturbance operator is of interest. In the deterministic case (if all A_0^i, B_0^i vanish) it is equal to the H^∞-norm of the associated rational transfer matrix. Thus $\|\boldsymbol{L}\|$ can be seen as a generalized H^∞-type norm for the stochastic system (1). To formulate a bounded real lemma for system (1) we define the affine linear operators

$$\begin{aligned} P(X) &= A^*X + XA + \sum_{i=1}^N A_0^{i\,*} X A_0^i - C^*C \\ S(X) &= XB + \sum_{i=1}^N A_0^{i\,*} X B_0^i - C^*D \\ Q(X) &= \sum_{i=1}^N B_0^{i\,*} X B_0^i + \gamma^2 I - D^*D, \quad \gamma > 0. \end{aligned}$$

Theorem 2.2 (Bounded Real Lemma, [EHP96]) *System (1) is internally stable and $\|\boldsymbol{L}\| < \gamma$ if and only if there exists $X < 0$ such that $Q(X) > 0$ and*

$$P(X) - S(X)Q(X)^{-1}S(X)^* \quad > \quad 0. \tag{2}$$

3 Resolvent positive operators

Let $\mathcal{H}^n \subset \mathbb{K}^{n\times n}, \mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ denote the real space of real or complex $n \times n$ Hermitian matrices, endowed with the Frobenius inner product $\langle X, Y\rangle = \operatorname{trace}(XY)$ and the corresponding (Frobenius) norm $\|\cdot\|$. By $\mathcal{H}^n_+ := \{X \in \mathcal{H}^n \mid X \geq 0\}$ we denote the closed convex cone of nonnegative definite matrices and by $\operatorname{int}(\mathcal{H}^n_+)$ its inner, i.e. the open cone of positive definite matrices. The cone $\mathcal{H}^n_+$ induces a partial ordering on $\mathcal{H}^n$: we write $X \geq Y$, if $X - Y \in \mathcal{H}^n_+$.

Proposition 3.1 *(i) The cone $\mathcal{H}^n_+$ is proper, i.e. it is closed, $\operatorname{int}(\mathcal{H}^n_+) \neq \emptyset$, and $\mathcal{H}^n_+ \cap -\mathcal{H}^n_+ = \{0\}$.*

(ii) The cone $\mathcal{H}^n_+$ is self-dual in $\mathcal{H}^n_+$, i.e. $\mathcal{H}^n_+ = (\mathcal{H}^n_+)^* := \{X \in \mathcal{H}^n \mid \forall Y \in \mathcal{H}^n_+ : \langle X, Y\rangle \geq 0\}$.

(iii) If $X, Y \in \mathcal{H}^n_+$ and $\langle X, Y\rangle = 0$, then $XY = YX = 0$.

Remark 3.2 Most of the results in this section can be transferred to the more general situation of an arbitrary finite-dimensional space ordered by a proper cone, but in view of our applications we restrict ourselves to the space $\mathcal{H}^n$ and the cone $\mathcal{H}^n_+$. For proofs and more details consult e.g. [KR50], [Sch65], [Els74], [Are87], [BP94], [DH99].

For a linear operator $\mathcal{T}$ let $\sigma(\mathcal{T})$ denote the spectrum, $\rho(\mathcal{T}) = \max\{|\lambda|; \lambda \in \sigma(\mathcal{T})\}$ the spectral radius, and $\beta(\mathcal{T}) = \max\{\operatorname{Re}(\lambda); \lambda \in \sigma(\mathcal{T})\}$ the spectral abscissa. The identity map is denoted by I, irrespective of the space it acts on.

Definition 3.3 *A linear operator $\mathcal{T} : \mathcal{H}^n \to \mathcal{H}^m$ is called* positive *($\mathcal{T} \geq 0$) if it maps $(\mathcal{H}^n_+)$ to $\mathcal{H}^m_+$. If $n = m$ then $\mathcal{T}$ is called* inverse positive *if $\mathcal{T}^{-1}$ exists and is positive, and* resolvent positive *if for all sufficiently large $\alpha > 0$ the operator $\alpha I - \mathcal{T}$ is inverse positive. For linear operators $\mathcal{S}, \mathcal{T} : \mathcal{H}^n \to \mathcal{H}^m$ we write $\mathcal{S} \geq \mathcal{T}$ if $\mathcal{S} - \mathcal{T}$ is positive.*

Remark 3.4 Since $\mathcal{H}^n_+$ is self-dual the adjoint operator $\mathcal{T}^*$ has the same positivity properties as $\mathcal{T}$.

Examples 3.5 (i) Let $A_0 \in \mathbb{K}^{n\times n}$, then the operator $\Pi : \mathcal{H}^n \to \mathcal{H}^n$ defined by $\Pi(X) := A_0^* X A_0$ is positive. If A_0 is nonsingular, then it is also inverse positive.
(ii) All positive operators Π are also resolvent positive, since for $\alpha > \rho(\Pi)$ the resolvent $(\alpha I - \Pi)^{-1} = \sum_{k=0}^\infty \alpha^{-(k+1)}\Pi^k$ is positive.
(iii) Given $A \in \mathbb{K}^{n\times n}$, the associated *Lyapunov operator* $\mathcal{L}_A : \mathcal{H}^n \to \mathcal{H}^n$, $\mathcal{L}_A(X) := A^*X + XA$, is resolvent positive but, in general, not positive. It is well known that $\mathcal{L}_A$ is inverse positive if and only if $\sigma(A) \subset \mathbb{C}_+$. Since $\alpha I - \mathcal{L}_A = \mathcal{L}_{\frac{\alpha}{2}I - A}$, the resolvent is positive for $\alpha > 2\beta(A)$.

Theorem 3.6 *Let $\mathcal{T} : \mathcal{H}^n \to \mathcal{H}^n$ be a resolvent positive linear operator. Then the following assertions hold:*

(i) $\beta(\mathcal{T}) \in \sigma(\mathcal{T})$ and there exists a matrix $V \in \mathcal{H}^n_+ \setminus \{0\}$ such that $\mathcal{T}(V) = \beta(\mathcal{T})V$.

(ii) If $\mathcal{S} : \mathcal{H}^n \to \mathcal{H}^n$ is a linear operator with $\mathcal{S} \geq \mathcal{T}$ then S is resolvent-positive and $\beta(\mathcal{S}) \geq \beta(\mathcal{T})$.

(iii) $\alpha I - \mathcal{T}$ is inverse positive $\iff \alpha > \beta(\mathcal{T})$.

Corollary 3.7 ([Sch65]) *Let $\mathcal{L} : \mathcal{H}^n \to \mathcal{H}^n$ be resolvent positive and $\Pi : \mathcal{H}^n \to \mathcal{H}^n$ be positive. Then the following are equivalent:*

(i) $\mathcal{L} + \Pi$ is stable, i.e. $\sigma(\mathcal{L} + \Pi) \subset \mathbb{C}_-$.

(ii) $-(\mathcal{L}+\Pi)$ *is inverse positive.*

(iii) $\sigma(\mathcal{L}) \subset \mathbb{C}_-$ *and* $\rho\left(\mathcal{L}^{-1}\Pi\right) < 1$.

(iv) $\exists X > 0: \quad (\mathcal{L}_A + \Pi)(X) < 0$.

Remark 3.8 ([Kha80]) If $\mathcal{L} = \mathcal{L}_A$ as in Example 3.5 (iii) and

$$\Pi(X) = \sum_{i=1}^{N} A_0^{(i)*} X A_0^{(i)}, \quad A_0^{(i)} \in \mathbb{K}^{n\times n},$$

then each of the conditions (i), (ii), and (iv) in Corollary 3.7 is equivalent to the exponential mean-square stability (compare Definition 2.1) of the linear Itô differential equation

$$dx_t \quad = \quad Ax(t)dt + \sum_{i=1}^{N} A_0^{(i)} x(t) dw_i(t).$$

Corollary 3.7 thus can be regarded as a generalization of Lyapunov's stability theorem for deterministic differential equations.

As in the latter context, we can weaken the definiteness conditions in Corollary 3.7 (iv), if (A, G) is observable:

Theorem 3.9 *Let* (A, G) *be observable,* $G \geq 0$*, and assume*

$$\exists X \leq 0 : \mathcal{L}_A(X) + \Pi(X) \geq G.$$

Then $X < 0$ *and* $\mathcal{L}_A + \Pi$ *is stable.*

4 A rational matrix operator

In order to analyze inequality (2) we introduce the rational matrix operator $\mathcal{R} : \mathcal{H}^n \to \mathcal{H}^n$, given by

$$\mathcal{R}(X) \quad = \quad P(X) - S(X)Q(X)^{-1}S(X)^*. \tag{3}$$

Here P, S, and Q are affine linear matrix operators from $\mathcal{H}^n$ to $\mathcal{H}^n$, $\mathbb{K}^{n\times\ell}$, and $\mathcal{H}^\ell$, respectively:

$$\begin{aligned} P(X) &= A^*X + XA + \Pi_1(X) + P_0 \\ S(X) &= XB + \Sigma(X) + S_0 \\ Q(X) &= \Pi_2(X) + Q_0, \end{aligned}$$

with $A \in \mathbb{K}^{n\times n}$, $P_0 \in \mathcal{H}^n$, $B, S_0 \in \mathbb{K}^{n\times\ell}$, $Q_0 \in \mathcal{H}^\ell$. We assume that the linear operators Π_1, Π_2, and Σ are the components of a *positive* linear operator

$$\Pi \quad : X \mapsto \begin{bmatrix} \Pi_1(X) & \Sigma(X) \\ \Sigma(X)^* & \Pi_2(X) \end{bmatrix} : \mathcal{H}^n \to \mathcal{H}^{n+\ell}. \tag{4}$$

In particular $\Pi_1 : \mathcal{H}^n \to \mathcal{H}^n$ and $\Pi_2 : \mathcal{H}^n \to \mathcal{H}^\ell$ are positive. We write

$$M \quad := \quad \begin{bmatrix} P_0 & S_0 \\ S_0^* & Q_0 \end{bmatrix} \in \mathcal{H}^{n+\ell}. \tag{5}$$

No restrictions on the inertia of M or any of its submatrices are imposed but we restrict the domain of $\mathcal{R}$ to

$$\operatorname{dom}\mathcal{R} \quad := \quad \{X \in \mathcal{H}^n \mid Q(X) > 0\}.$$

If $\operatorname{dom}\mathcal{R}$ happens to be empty, the further discussion is void. But note that $X \in \operatorname{dom}\mathcal{R}$ implies $Y \in \operatorname{dom}\mathcal{R}$ for all $Y \geq X$. We are interested in the rational matrix equation

$$\mathcal{R}(X) \quad = \quad 0, \tag{6}$$

and the corresponding strict and nonstrict inequalities.

Remark 4.1 *Equation (6) reduces to the CARE*

$$A^*X + XA + P_0 - (XB + S_0)\, Q_0^{-1} (XB + S_0)^* \quad = \quad 0$$

if $\Pi = 0$*, and to the DARE*

$$\begin{aligned} X \quad &= \quad {A_0}^* X A_0 + P_0 \\ &- \left({A_0}^* X B_0 + S_0\right)\left({B_0}^* X B_0 + Q_0\right)^{-1}\left({A_0}^* X B_0 + S_0\right)^*, \end{aligned}$$

if $A = -\frac{1}{2}I$, $B = 0$, *and* $\Pi_1(X) = {A_0}^* X A_0$, $\Pi_2(X) = {B_0}^* X B_0$, *and* $\Sigma(X) = {A_0}^* X B_0$.
The definite version of equation (6) with $\Sigma = 0$ *and* $M > 0$ *was first studied by Wonham in [Won68].*

In order to apply Newton's method to equation (6) we need to study the Taylor expansion of $\mathcal{R}$. We first calculate the Fréchet derivative $\mathcal{R}'_X$ of $\mathcal{R}$ at points $X \in \operatorname{dom}\mathcal{R}$. We write P_X, S_X, and Q_X instead of $P(X)$, $S(X)$, and $Q(X)$. The following proposition can be verified by a direct calculation, see [DH99].

Proposition 4.2 *Set* $A_X := A - BQ_X^{-1}S_X^*$ *and*

$$\Pi_X(H) \quad := \quad \begin{bmatrix} I \\ -Q_X^{-1}S_X^* \end{bmatrix}^* \Pi(H) \begin{bmatrix} I \\ -Q_X^{-1}S_X^* \end{bmatrix}.$$

Then $\mathcal{R}'_X \;=\; \mathcal{L}_{A_X} + \Pi_X$.

Remark 4.3 By Remark 3.4, Example 3.5 (iii) and Theorem 3.6 (ii) both $\mathcal{R}'_X$ and $(\mathcal{R}'_X)^*$ are resolvent positive. The adjoint operator $(\mathcal{R}'_X)^*$ is given by

$$\begin{aligned} H \quad \mapsto \quad & (A - BQ_X^{-1}S_X^*)H + H(A - BQ_X^{-1}S_X^*)^* \\ & + \Pi_1(H) - \Sigma\left(Q_X^{-1}S_X^*H\right) - \Sigma\left(HS_XQ_X^{-1}\right) \\ & + \Pi_2\left(Q_X^{-1}S_X^*HS_XQ_X^{-1}\right). \end{aligned}$$

Thus $Q_X^{-1}S_X^*H \;=\; Q_Y^{-1}S_Y^*H$ implies $(\mathcal{R}'_X)^*(H) = (\mathcal{R}'_Y)^*(H)$.

We now introduce an appropriate stabilizability concept for the matrix operator $\mathcal{R}$:

Definition 4.4 *A matrix* $X \in \operatorname{dom}\mathcal{R}$ *is called* stabilizing for $\mathcal{R}$ *if* $X \in \operatorname{dom}\mathcal{R}$ *and* $\mathcal{R}'_X \subset \mathbb{C}_- = \{s \in \mathbb{C} \mid \operatorname{Re} s < 0\}$ *and* almost stabilizing *if* $\mathcal{R}'_X \subset \mathbb{C}_- \cup i\mathbb{R}$. *We call* $\mathcal{R}$ (almost) stabilizable *if there exists an (almost) stabilizing matrix* X *for* $\mathcal{R}$.

Remark 4.5 (i) A sufficient condition for a matrix X to stabilize $\mathcal{R}$ is that A_X is stable and

$$I \;>\; \mathcal{L}_{A_X}^{-1}\left(\Pi_X(I)\right) = \int_0^\infty e^{tA_X^*}\Pi_X(I)e^{tA_X}\,dt.$$

This condition was given in [Won68]. In [Hau72] examples were provided, that it is not a necessary condition.
(ii) In the special case that $\mathcal{R}(X) = 0$ is the CARE, a matrix X is stabilizing, if and only if all eigenvalues of $A - BQ_0^{-1}(S_0^* + B^*X)$ lie in the open left half plane.
(iii) In the special case that $\mathcal{R}(X) = 0$ is the DARE, a matrix X is stabilizing, if and only if all eigenvalues of $A_0 - B_0\left(Q_0 + B_0^*XB_0\right)^{-1}(S_0^* + B_0^*XA_0)$ lie in the open unit disk.

It is convenient to write for $Y, Z \in \operatorname{dom}\mathcal{R}$

$$M_Y := \begin{bmatrix} I \\ -Q_Y^{-1}S_Y^* \end{bmatrix}^* M \begin{bmatrix} I \\ -Q_Y^{-1}S_Y^* \end{bmatrix} \tag{7}$$

and

$$\Phi_Z(Y) \;:=\; \left(S_YQ_Y^{-1} - S_ZQ_Z^{-1}\right)Q_Y\left(Q_Y^{-1}S_Y^* - Q_Z^{-1}S_Z^*\right).$$

Note that $\Phi_Z(Y) \geq 0$ and $\Phi_Y(Y) = 0$.
The following proposition shows that $-\Phi_Z(Y)$ is the second order remainder term of the Taylor expansion of $\mathcal{R}$ at Z in the direction $Y - Z$. In particular, the positivity of $\Phi_Z(Y)$ implies that $\mathcal{R}$ is a *concave* operator on $\operatorname{dom} R$ in the sense that its values lie below the values of its tangents at any point $Z \in \operatorname{dom}\mathcal{R}$ (with respect to the order of $\mathcal{H}^n$) : $\mathcal{R}(Y) \leq \mathcal{R}(Z) + \mathcal{R}'_Z(Y - Z)$ for all $Y, Z \in \operatorname{dom}\mathcal{R}$.

Proposition 4.6 *Let $Y, Z \in \operatorname{dom}\mathcal{R}$. Then the following identities hold:*

(i) $\mathcal{R}(Y) = \mathcal{R}'_Z(Y) + M_Z - \Phi_Z(Y)$.

(ii) $\mathcal{R}(Z) + R'_Z(Y - Z) = \mathcal{R}(Y) + \Phi_Z(Y) \geq \mathcal{R}(Y)$.
In particular $\mathcal{R}$ is a concave map on $\operatorname{dom}\mathcal{R}$.

A proof of these basic properties can be found in [DH99].

5 Newton's method

In this section we derive our main result. We present an iterative algorithm to solve the rational matrix equation $\mathcal{R}(X) = 0$. It is a non-local version of Newton's method, which is already well-established in the cases of the deterministic and the definite algebraic Riccati equation; we mention [Kle68], [Won68], [Cop74], [Meh91], [LR95].
The method can be applied under the conditions, that $\mathcal{R}$ is stabilizable and that the inequality $\mathcal{R}(X) \geq 0$ is solvable. Under these conditions it will be shown that convergence takes place if the algorithm starts at any stabilizing initial matrix X_0 (see Definition 4.4).

Using Proposition 4.6 (i) we can write the standard Newton-iteration for our problem in the following form:

$$\begin{aligned} X_{k+1} &= X_k - \left(\mathcal{R}'_{X_k}\right)^{-1}\left(\mathcal{R}(X_k)\right) \\ &= -\left(\mathcal{R}'_{X_k}\right)^{-1}\left(M_{X_k}\right), \end{aligned} \tag{8}$$

where $\mathcal{R}'_{X_k}$ is known from Proposition 4.2, and M_X was defined in (7). In each iteration step the following n^2-dimensional linear system must be solved in order to obtain X_{k+1}: $A_{X_k}^*X + XA_{X_k} + \Pi_{X_k}(X) = M_{X_k}$.

Theorem 5.1 *Assume that there exist a solution $\hat{X} \in \operatorname{dom}\mathcal{R}$ of $\mathcal{R}(X) \geq 0$ and a stabilizing matrix X_0 for $\mathcal{R}$. Then the iteration scheme (8) defines a sequence (X_k) in $\operatorname{dom}\mathcal{R}$ with the following properties:*

(i) $\forall k = 1, 2, \ldots :\ X_k \geq X_{k+1} \geq \hat{X}$ *and* $\mathcal{R}(X_k) \leq 0$.

(ii) $\forall k = 0, 1, 2, \ldots :\ \mathcal{R}'_{X_k}$ *is stable.*

(iii) (X_k) *converges to a limit matrix* $X_\infty \in \operatorname{dom}\mathcal{R}$.

(iv) $\mathcal{R}(X_\infty) = 0$, X_∞ *is the largest solution of* $\mathcal{R}(X) \geq 0$ *and* $\sigma\left(\mathcal{R}'_{X_\infty}\right) \subset \mathbb{C}_- \cup i\mathbb{R}$.

Proof: We prove (i) and (ii) inductively.
By assumption $\mathcal{R}'_{X_0}$ is stable, which settles the case $k = 0$. Suppose that $X_0, \ldots, X_k$ have been constructed such that $\mathcal{R}'_{X_i}$ is stable for $i = 0, \ldots, k$, $X_1 \geq \ldots \geq X_k$ and $\mathcal{R}_{X_i} \leq 0$ for $i = 1, \ldots, k$. Then X_{k+1} is well defined by (8) and satisfies

$$\mathcal{R}'_{X_k}(X_k - X_{k+1}) \;=\; \mathcal{R}(X_k). \tag{9}$$

We first prove $X_{k+1} \geq \hat{X}$. By the concavity of $\mathcal{R}$ (Proposition 4.6 (ii)) we have

$$\begin{aligned} \mathcal{R}'_{X_k}(\hat{X} - X_{k+1}) &= \mathcal{R}'_{X_k}(\hat{X} - X_k) + \mathcal{R}'_{X_k}(X_k - X_{k+1}) \\ &= \mathcal{R}'_{X_k}(\hat{X} - X_k) + \mathcal{R}(X_k) \overset{4.6\,(ii)}{\geq} \mathcal{R}(\hat{X}) \;\geq\; 0. \end{aligned} \tag{10}$$

Since $\mathcal{R}'_{X_k}$ is stable, i.e. $-\mathcal{R}'_{X_k}$ is inverse positive by Corollary 3.7, we have $\hat{X} \leq X_{k+1}$ and hence $X_{k+1} \in \operatorname{dom}\mathcal{R}$.
By the same argument, if $k \geq 1$ it follows directly from (9) and $\mathcal{R}(X_k) \leq 0$ that $X_k - X_{k+1} \geq 0$. It remains to show that $\mathcal{R}'_{X_{k+1}}$ is stable and $\mathcal{R}(X_{k+1}) \leq 0$. Again by the concavity of $\mathcal{R}$ we have

$$0 \;\leq\; \mathcal{R}(\hat{X}) \;\leq\; \mathcal{R}(X_{k+1}) + \mathcal{R}'_{X_{k+1}}(\hat{X} - X_{k+1}). \tag{11}$$

From 4.6 (i) and (8) we have

$$\mathcal{R}(X_{k+1}) \;=\; -\Phi_{X_k}(X_{k+1}) \;\leq\; 0, \tag{12}$$

which proves $\mathcal{R}(X_{k+1}) \leq 0$, and together with (11) we obtain

$$\mathcal{R}'_{X_{k+1}}(\hat{X} - X_{k+1}) \;\geq\; \Phi_{X_k}(X_{k+1}) \;\geq\; 0. \tag{13}$$

Now let us assume that $\mathcal{R}'_{X_{k+1}}$ is not stable. By Remark 4.3 both $\mathcal{R}'_{X_{k+1}}$ and its adjoint are resolvent positive; moreover they have the same spectral abscissa. Thus by Theorem 3.6 the instability of $\mathcal{R}'_{X_{k+1}}$ is equivalent to the following condition:

$$\exists V \in \mathcal{H}^n_+ \setminus \{0\}, \lambda \geq 0 : \qquad {\mathcal{R}'}^*_{X_{k+1}}(V) = \lambda V. \quad (14)$$

On the one hand this implies

$$\left\langle V \,,\, \mathcal{R}'_{X_{k+1}}(\hat{X} - X_{k+1}) \right\rangle \;=\; \left\langle \lambda V \,,\, \hat{X} - X_{k+1} \right\rangle \;\leq\; 0.$$

On the other hand we have from (13)

$$\left\langle V \,,\, \mathcal{R}'_{X_{k+1}}(\hat{X} - X_{k+1}) \right\rangle \;\geq\; \langle V \,,\, \Phi_{X_k}(X_{k+1}) \rangle \;\geq\; 0.$$

Combined with the previous inequality this gives

$$\langle V \,,\, \Phi_{X_k}(X_{k+1}) \rangle = 0 \quad \text{i.e.} \quad \Phi_{X_k}(X_{k+1}) V = 0,$$

where we have used Proposition 3.1. Thus by the definition of Φ: $Q^{-1}_{X_{k+1}} S^*_{X_{k+1}} V = Q^{-1}_{X_k} S^*_{X_k} V$.
By Remark 4.3 this implies $\lambda V = {\mathcal{R}'}^*_{X_{k+1}}(V) = {\mathcal{R}'}^*_{X_k}(V)$, contradicting the stability of $\mathcal{R}'_{X_k}$. Thus, our assumption was wrong, and $\mathcal{R}'_{X_{k+1}}$ is stable. This concludes our proof of (i) and (ii) by induction.
(iii) and (iv) follow easily from (i) and (ii). □

From Theorem 5.1 we infer an existence theorem for the equation $\mathcal{R}(X) = 0$, which generalizes existence theorems for the definite equations from LQ control theory. In [Won68] this result was given under the additional assumptions $\Sigma = 0$ and (A, P_0) detectable:

Corollary 5.2 *Assume $Q_0 > 0$, $P_0 \geq S_0^* Q_0^{-1} S_0$, and $\mathcal{R}$ is stabilizable. Then the equation $\mathcal{R}(X) = 0$ has a solution $X \geq 0$.*

Proof: The assumptions guarantee that $0 \in \operatorname{dom} \mathcal{R}$ and $\mathcal{R}(0) \geq 0$. Thus Theorem 5.1 can be applied with $\hat{X} = 0$ and a stabilizing X_0. □

To apply the algorithm proposed in Theorem 5.1 it is crucial and in many cases not trivial to find a stabilizing matrix X_0. An important special case, however, is when $0 \in \operatorname{dom} \mathcal{R}$ (i.e. $Q_0 > 0$) and $\mathcal{R}'_0$ is stable:

Theorem 5.3 *Let the pair (A, P_0) be observable, $P_0 \leq 0$, $Q_0 > 0$ and assume that the inequality $\mathcal{R}(X) \geq 0$ has a solution in $\operatorname{dom} \mathcal{R}$. Then the following are equivalent:*

(i) $\mathcal{R}'_0$ is stable.

(ii) $\exists X \in \operatorname{dom} \mathcal{R} :\ \mathcal{R}(X) \geq 0$ and $X \leq 0$.

(iii) $X \in \operatorname{dom} \mathcal{R}$ and $\mathcal{R}(X) \geq 0 \Rightarrow X \leq 0$.

(iv) $X \in \operatorname{dom} \mathcal{R}$ and $\mathcal{R}(X) \geq 0 \Rightarrow X < 0$.

The implication (i)⇒(iii) (as well as the trivial chain (iv)⇒(iii)⇒(ii)) also holds without the observability assumption.

Proof: (i)⇒(iii): By concavity

$$\mathcal{R}(X) \;\leq\; \mathcal{R}(0) + \mathcal{R}'_0(X) \;=\; P_0 - S_0 Q_0^{-1} S_0^* + \mathcal{R}'_0(X).$$

Thus $\mathcal{R}(X) \geq 0$ implies $\mathcal{R}'_0(X) \geq -P_0$ which by the stability of $\mathcal{R}'_0$ in turn implies $X \leq 0$.
The remaining non-trivial implications (ii)⇒(i) and (iii)⇒(iv) follow immediately from Theorem 3.9. □

6 Stabilizing solutions

As we have seen, the Newton iteration produces an almost stabilizing solution. In many cases it is interesting to know, if there exists a stabilizing solution. For instance we will see below, that this guarantees quadratic convergence of the Newton iteration.

Lemma 6.1 *Let $Y, Z \in \operatorname{dom} \mathcal{R}$ such that $\mathcal{R}(Y) \leq \mathcal{R}(Z)$, and $\sigma(\mathcal{R}'_Y) \subset \mathbb{C}_-$. Then $Y \geq Z$.*

Theorem 6.2 *The following are equivalent:*

(i) $\mathcal{R}$ is stabilizable and $\exists \hat{X} \in \operatorname{dom} \mathcal{R} :\ \mathcal{R}(\hat{X}) > 0$.

(ii) There exists a stabilizing solution of the equation $\mathcal{R}(X) = 0$.

Moreover, a stabilizing solution of the equation $\mathcal{R}(X) = 0$ is necessarily the largest solution of the inequality $\mathcal{R}(X) \geq 0$ and thus unique.

Proof: (i)⇒(ii): Let the sequence (X_k) be defined as in Theorem 5.1. Since $\mathcal{R}(\hat{X}) > 0$, inequality (10) holds in its strict form. Passing to the limit $k \to \infty$ yields

$$\mathcal{R}'_{X_\infty}(\hat{X} - X_\infty) \;>\; 0.$$

By continuity $\mathcal{R}'_{X_\infty}$ maps a whole neighbourhood of $\hat{X} - X_\infty \leq 0$ to $\operatorname{int}(\mathcal{H}^n_+)$. Thus $\mathcal{R}'_{X_\infty}$ is stable by Corollary 3.7, i.e. X_∞ is a stabilizing solution of the equation $\mathcal{R}(X) = 0$. If $\tilde{X}$ is another stabilizing solution of the equation $\mathcal{R}(X) = 0$, then by Lemma 6.1 $\tilde{X} \geq X_\infty$ and $X_\infty \geq \tilde{X}$, i.e. $X_\infty = \tilde{X}$.
(ii)⇒(i): Let X_∞ be stabilizing. Then $\mathcal{R}'_{X_\infty}$ is a regular operator and by the implicit function theorem the equation $\mathcal{R}(X) - \epsilon I = 0$ is solvable for sufficiently small ϵ in a neighbourhood of X_∞ in $\mathcal{H}^n$. Hence (i). □

If there exists a stabilizing solution then convergence of the sequence (X_k) defined in Theorem 5.1 is quadratic (compare [Meh91] for CARE and DARE).

Theorem 6.3 *Assume that there exists a stabilizing solution $X \in \operatorname{dom} \mathcal{R}$ to $\mathcal{R}(X) = 0$.*
Let the sequence (X_k) and its limit X_∞ be given as in Theorem 5.1. Then there exists a constant κ such that

$$\|X_{k+1} - X_\infty\| \quad \leq \quad \kappa \|X_k - X_\infty\|^2 .$$

For a proof of this theorem see [DH99].

Remark 6.4 The main results of the Sections 5 and 6 can be transferred e.g. to the more general situation, where $\mathcal{R} : \mathcal{X} \to \mathcal{X}$ is a concave operator in an arbitrary finite-dimensional space $\mathcal{X}$ ordered by a proper cone $\mathcal{C}$, such that $\mathcal{R}'_X$ is resolvent positive for all $X \in \operatorname{dom} \mathcal{R}$ and $\operatorname{dom} \mathcal{R} = \operatorname{dom} \mathcal{R} + \mathcal{C}$; see also [DH99].

7 Monotonicity and concavity

In this section we compare the largest solutions of the matrix inequalities $\mathcal{R}_0(X) \geq 0$, $\mathcal{R}_1(X) \geq 0$, where $\mathcal{R}_0$ and $\mathcal{R}_1$ are of the type (3); (for special cases see e.g. [Wim85], [RV88], [LR95]). It is convenient to do this in the framework of linear matrix inequalities:

Proposition 7.1 *Let* $\Lambda(X) := \begin{bmatrix} A^*X + XA & XB \\ B^*X & 0 \end{bmatrix}$ *and Π and M as in (5) and (4). Then for $X \in \operatorname{dom} \mathcal{R}$:*

(i) $\mathcal{R}(X) \geq 0 \iff \Pi(X) + \Lambda(X) + M \geq 0.$

(ii) $\mathcal{R}(X) > 0 \iff \Pi(X) + \Lambda(X) + M > 0.$

We will say that $\mathcal{R}$ is *the rational operator associated to* Π, Λ, and M.
For $i = 0, 1$ let $M_i \in \mathcal{H}^{n+\ell}$ and $\mathcal{R}_i$ be the rational matrix operators associated to Π, Λ, and M_i.

Theorem 7.2 *Assume that there exists a stabilizing solution X_0 to $\mathcal{R}_0(X) = 0$. Assume further that $\mathcal{R}_1$ is stabilizable and $M_1 \geq M_0$. Then there exists a stabilizing solution X_1 to $\mathcal{R}_1(X) = 0$ and $X_1 \geq X_0$.*

Proof: By Theorem 6.2 there exists an $\tilde{X} \in \operatorname{dom} \mathcal{R}_0$ such that $\mathcal{R}_0(\tilde{X}) > 0$. By $M_1 \geq M_0$ and Proposition 7.1 we have $\operatorname{dom} \mathcal{R}_0 \subset \operatorname{dom} \mathcal{R}_1$ and also $\mathcal{R}_1(\tilde{X}) > 0$. Thus again Theorem 6.2 gives us a unique stabilizing solution X_1 of $\mathcal{R}_1(X) = 0$. Another application of Proposition 7.1 yields $\mathcal{R}_1(X_0) \geq 0$. By Lemma 6.1, however, X_1 is the largest solution of the latter inequality, whence $X_1 \geq X_0$. □

An analogous argument shows that X_∞ depends on M in a concave fashion.

Theorem 7.3 *Let $M_0, M_1 \in \mathcal{H}^{n+\ell}$ be arbitrary, and assume that for $i = 0, 1$ there exist stabilizing solutions X_i of $\mathcal{R}_i(X) = 0$. For $t \in [0, 1]$ set $M_t := (1-t)M_0 + tM_1$ and denote the rational operator associated to Π, Λ, and M_t by $\mathcal{R}_t$. Then for all $t \in [0, 1]$ there exists a stabilizing solution X_t of $\mathcal{R}_t(X) = 0$ and $X_t \geq (1-t)X_0 + tX_1$.*

References

[Are87] W. Arendt. Resolvent positive operators. *Proc. London Math. Soc.*, 54(3):321–349, 1987.

[BP94] A. Berman and R. J. Plemmons. *Nonnegative Matrices in the Mathematical Sciences.* Classics in Applied Mathematics. SIAM, 1994.

[Cop74] W. A. Coppel. Matrix quadratic equations. *Bull. Amer. Math. Soc.*, 10:377–401, 1974.

[DH99] T. Damm and D. Hinrichsen. Newton's method for a rational matrix equation occuring in stochastic control. Report 443, Institut für Dynamische Systeme, Universität Bremen, 1999.

[EHP96] A. El Bouhtouri, D. Hinrichsen, and A. J. Pritchard. On the disturbance attenuation problem for a wide class of time invariant linear stochastic systems. *to appear in: Stochastics and Stochastics Reports*, 1996.

[EHP98] A. El Bouhtouri, D. Hinrichsen, and A. J. Pritchard. H^∞ type control for discrete-time stochastic systems. Report 424, Institut für Dynamische Systeme, Universität Bremen, January 1998.

[Els74] L. Elsner. Quasimonotonie und Ungleichungen in halbgeordneten Räumen. *Lin. Alg. Appl.*, 8:249–261, 1974.

[Hau72] U. G. Haussmann. Stabilization of linear systems with multiplicative noise. In Ruth F. Curtain, editor, *Stability of Stochastic Dynamical Systems*, volume 294 of *Lecture Notes in Mathematics*, pages 125–130. Springer-Verlag, July 1972.

[Kha80] R. Z. Khasminskij. *Stochastic Stability of Differential Equations.* Sijthoff & Noordhoff, Alphen aan den Rijn, NL, 1980.

[Kle68] D. L. Kleinman. On an iterative technique for Riccati equation computation. *IEEE Trans. Automat. Contr.*, AC-13:114–115, 1968.

[KR50] M. G. Krein and M. A. Rutman. Linear operators leaving invariant a cone in a Banach space. *Amer. Math. Soc. Transl.*, 26:199–325, 1950.

[LR95] P. Lancaster and L. Rodman. *Algebraic Riccati Equations.* Oxford, 1995.

[Meh91] V. Mehrmann. *The Autonomous Linear Quadratic Control Problem, Theory and Numerical Solution.* Number 163 in Lecture Notes in Control and Information Sciences. Springer-Verlag, Heidelberg, July 1991.

[RV88] A. C. M. Ran and R. Vreugdenhil. Existence and comparison theorems for algebraic Riccati equations for continuous- and discrete-time systems. *Lin. Alg. Appl.*, 99:63–83, 1988.

[Sch65] H. Schneider. Positive operators and an inertia theorem. *Numerische Mathematik*, 7:11–17, 1965.

[Wim85] H. K. Wimmer. Monotonicity of maximal solutions of algebraic Riccati equations. *Systems & Control Letters*, 5:317–319, 1985.

[Won68] W. M. Wonham. On a matrix Riccati equation of stochastic control. *SIAM J. Control Optim.*, 6:681–698, 1968.