Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Positive invariance tests with efficient Hessian matrix eigenvalues bounds

M. Mönnigmann *

* *Technische Universität Braunschweig, Institute for Heat and Fuel Technology, P.O. Box 3329, 38023 Braunschweig, Germany (email: m.moennigmann@tu-bs.de)*

Abstract: We investigate two simple sufficient criteria for positive invariance of sets in the domain of n-dimensional nonlinear autonomous discrete time systems. These criteria are derived from the exact Taylor expansion with linear and quadratic remainder terms. By a simple example we demonstrate that systems exist for which positive invariance can be established with the second order criterion but not with the first order criterion. Since the second order criterion requires the Hessian matrices of the model equations, this criterion is computationally expensive. We show, however, that the second order criterion can be evaluated at a surprisingly low computational cost. Specifically, we show that the computational complexity is an order of magnitude lower than the calculation of the Hessian matrices.

1. INTRODUCTION

The notion of positively invariant (p.i.) sets plays an important role in many problems in control theory (Blanchini, 1999). In the present paper we investigate positive invariance of sets for nonlinear autonomous discrete time systems. More precisely, let $U \subset R^n$ be open and let $f : U \to R^n$ be a twice continuously differentiable function. We are interested in establishing positive invariance of balls around a locally stable fixpoint x_0 of systems of the form

$$x(t_{k+1}) = f(x(t_k)), \quad x(0) = x_0, \tag{1}$$

if such a ball exist. In particular we are interested in computational methods for establishing positive invariance that can be carried out automatically. We present such a method and discuss its computational complexity.

First we sketch the reason for using Hessian matrices of the f_i, i.e. we explain why it is of interest to use a second order approximation of the r.h.s. of f in (1). Note the following reasoning is informal at times. Section 2 treats the same idea in a more formal fashion.

Let $Q \subset U$ be convex and let $\Delta x := x - x_0$. According to the mean value theorem or Taylor's theorem we know that for any $x \in Q$ and any $x_0 \in Q$

$$f(x) = f(x_0) + S(x, x_0) \tag{2}$$

where the remainder to linear order, $S(x, x_0)$, can be calculated (Heuser, 1991) from

$$S_i(x, x_0) = \sum_{\rho=1}^{n} \Delta x_\rho \int_0^1 \frac{\partial f_i(x_0 + s\Delta x)}{\partial x_\rho}\, ds, \tag{3}$$

for $i = 1, \ldots, n$. Similarly, according to Taylor's theorem

$$f(x) = f(x_0) + \nabla f(x_0)\, \Delta x + T(x, x_0) \tag{4}$$

where the remainder to quadratic order, $T(x, x_0)$, can be calculated (Heuser, 1991) from

$$T_i(x, x_0) = \sum_{\rho,\sigma=1}^{n} \Delta x_\rho \Delta x_\sigma \int_0^1 \frac{\partial^2 f_i(x_0 + s\Delta x)}{\partial x_\rho \partial x_\sigma}\, (1 - s)\, ds, \tag{5}$$

for $i = 1, \ldots, n$. By assumption of local stability of system (1) the mapping is globally stable

$$x \to f(x_0) + \nabla f(x_0) \Delta x \tag{6}$$

since $\nabla f(x_0)$ has only eigenvalues in the unit circle. When using the Taylor theorem with remainder to linear order, however, $\nabla f(x_0)\, \Delta x$ in the mapping (6) is replaced by $S(x, x_0)$. In order to evaluate $S(x, x_0)$, the gradient $\nabla f(x)$ needs to be evaluated not only at x_0, but also away from x_0 according to (3). Therefore information on the eigenvalues of $\nabla f(x)$ at x_0 only is no longer sufficient to infer stability. When using the second order approximation (4), in contrast, the globally stable linear order is retained, cf. the second term in the r.h.s. of (4), and information on f away from x_0 is only needed to second order. Loosely speaking, the approximation to second order (4) therefore makes better use of the fact that the system (1) is known to be locally stable at x_0. This idea is treated in more detail in section 2.

Since the calculation of bounds to second order involves second derivatives of $f(x)$, the second order criterion must be expected to be computationally more expensive than its first order counterpart. In section 3 we show, however, that the necessary second order information can be obtained without ever calculating the Hessian matrices of the f_i. More generally, section 3 summarizes recent results on the complexity of Hessian matrix eigenvalue bound calculation that are likely to be useful beyond the purpose they serve here.

Section 4 gives a simple illustrative example for which positive invariance on a set can be established with the second order remainder but not with the first order remainder. Conclusions are given at the end of the paper.

2. SIMPLE SUFFICIENT CONDITIONS FOR POSITIVE INVARIANCE ON BALLS

We call a set $M \subset U$ positively invariant (p.i.) for system (1) if for all $x(t_0) = x_0 \in M$ the system stays in M for all t, i.e. $x(t_k) \in M$ for all $k > 0$. There is a large body of literature on positive invariance on polytopes (see Blanchini (1999) and references therein). In what follows we restrict ourselves to the simpler case of positive invariance on balls.

Using bounds on the eigenvalues of the Jacobian of f or on the eigenvalues of the Hessians of the f_i, $i = 1, \ldots, n$ we can state criteria for positive invariance of balls around a fixpoint. This is clarified in the following simple propositions 1 and 2. Let $||\cdot||_2$ denote the 2-norm on R^n. By $B(\rho)$ denote the closed ball $B(\rho) = \{x \in R^n : ||x||_2 \leq \rho\}$ with radius $\rho > 0$. Furthermore, for any real symmetric matrix M denote the largest eigenvalue of M by $\lambda_{\max}(M)$.

Proposition 1. (suff. cond. from linear remainder S) Let the assumptions be as for system (1). If there exists a $\rho > 0$ such that for

$$\sigma := \max_{x \in B(\rho)} \sqrt{\lambda_{\max}\left((\nabla f(x))^T \nabla f(x)\right)} \tag{7}$$

the inequality

$$||f(x_0)||_2 + \sigma\, ||\Delta x||_2 \leq \rho \tag{8}$$

holds for all $x \in B(\rho)$, then $B(\rho)$ is p.i. for the dynamical system (1).

Proof. We rewrite $S(x, x_0)$ as defined in (3) as $S(x, x_0) = A(x_0 + s\Delta x)\, \Delta x$, where

$$A_{ij}(x_0, x) := \int_0^1 \frac{\partial f_i(x_0 + s\Delta x)}{\partial x_j} ds,$$

for $i, j = 1, \ldots, n$. Then

$$||S(x, x_0)||_2 \leq ||A(x_0 + s\Delta x)||_2\, ||\Delta x||_2 \tag{9}$$

where the vector and induced matrix 2-norm are used. Furthermore

$$\begin{aligned} ||A(x_0 + s\Delta x)||_2 &\leq \left(\lambda_{\max}\left(A^T(x_0 + s\Delta x) A(x_0 + s\Delta x)\right)\right)^{1/2} \\ &\leq \sigma \end{aligned} \tag{10}$$

where the first inequality is a standard result from linear algebra and the second inequality holds by definition of σ. By applying the 2-norm and the triangle inequality to (2) and by combining (9) and (10) we get

$$\begin{aligned} ||f(x)||_2 &\leq ||f(x_0)||_2 + ||S(x, x_0)||_2 \\ &\leq ||f(x_0)||_2 + \sigma\, ||\Delta x||_2 \end{aligned} \tag{11}$$

for all x in $B(\rho)$. Combining (11) with (8) we get $||f(x)||_2 \leq \rho$ for all $x \in B(\rho)$. Since this implies $||x_{k+1}||_2 = ||f(x_k)||_2 \leq \rho$ for all $x_k \in B(\rho)$, trajectories that enter the ball $B(\rho)$ never leave it. Therefore $B(\rho)$ is p.i. for the dynamical system (1).

Similarly, we can state a sufficient condition for positive invariance based on the quadratic remainder (5).

Proposition 2. (suff. cond. from quadratic remainder T) If there exists a $\rho > 0$ such that for

$$\underline{\lambda}_i := \min_{x \in B(\rho)} \lambda_{\min}\left(H^{(i)}(x)\right), \quad i = 1, \ldots, n, \tag{12}$$

$$\bar{\lambda}_i := \max_{x \in B(\rho)} \lambda_{\max}\left(H^{(i)}(x)\right), \quad i = 1, \ldots, n, \tag{13}$$

and

$$\tau := \left(\sum_{i=1}^{n} \max\left(\underline{\lambda}_i^2, \bar{\lambda}_i^2\right)\right)^{1/2} \tag{14}$$

the inequality

$$||f(x_0)||_2 + ||\nabla f(x_0)||_2\, ||\Delta x||_2 + \frac{1}{2}\tau\, ||\Delta x||_2^2 \leq \rho \tag{15}$$

holds for all $x \in B(\rho)$, then $B(\rho)$ is p.i. for the dynamical system (1).

We use the following lemma to prove proposition 2.

Lemma 3. Let the assumptions be as for system (1) and let the second order remainder T be defined as in (5). Then for all $i = 1, \ldots, n$ and for all x and x_0 in $B(\rho)$

$$\frac{1}{2}\,\underline{\lambda}_i ||\Delta x||_2^2 \leq T_i(x, x_0) \leq \frac{1}{2}\,\bar{\lambda}_i ||\Delta x||_2^2, \tag{16}$$

where $\underline{\lambda}_i$ and $\bar{\lambda}_i$ are defined as in (12) and (13), respectively.

Proof. Let $x \in B(\rho)$, $x_0 \in B(\rho)$, $s \in [0, 1]$, and $i \in \{1, \ldots, n\}$ be arbitrary but fixed. Since $B(\rho)$ is convex, all points on the line $x_0 + s\,\Delta x$ lie in $B(\rho)$. According to a standard result on quadratic forms we have, for all $\xi \in R^n$ and all $s \in [0, 1]$,

$$\sum_{\rho,\sigma=1}^{n} \xi_\rho H^{(i)}(x_0 + s\Delta x)\xi_\sigma \leq ||\xi||_2^2\, \lambda_{\max}(H^{(i)}(x_0 + s\Delta x)). \tag{17}$$

Furthermore for all $s \in [0, 1]$

$$\begin{aligned} \lambda_{\max}(H^{(i)}(x_0 + s\Delta x)) &\leq \max_{\substack{x \in B(\rho) \\ x_0 \in B(\rho) \\ s \in [0,1]}} \lambda_{\max}(H^{(i)}(x_0 + s\Delta x)) \\ &= \max_{x \in B(\rho)} \lambda_{\max}(H^{(i)}(x)), \end{aligned} \tag{18}$$

where the equality in the last step holds because the set $V = \{x_0 + s\Delta x : x \in B(\rho), x_0 \in B(\rho), s \in [0, 1]\}$ is the set of all points on all lines between arbitrary points x_0 and x in $B(\rho)$. For this set we have, however, $V = B(\rho)$. Since (17) holds for all $\xi \in R$, it holds for $\xi = \Delta x = x - x_0$ in particular. Combining (17) and (18) therefore yields the upper bound in (16). The lower bound can be proved in the same fashion.

Using lemma 3 we can now easily prove proposition 2.

Proof of prop. 2. From (16) we infer

$$0 \leq T_i^2(x, x_0) \leq \left(\frac{1}{2}||\Delta x||_2^2\right)^2 \max\left(\underline{\lambda}_i^2, \bar{\lambda}_i^2\right).$$

By summing over $i = 1, \ldots, n$ and taking the square root we get

$$0 \leq ||T(x, x_0)||_2 \leq \frac{1}{2}||\Delta x||_2^2 \left(\sum_{i=1}^{n} \max(\underline{\lambda}_i^2, \bar{\lambda}_i^2)\right)^{1/2} \tag{19}$$

On the other hand, applying the norm $||\cdot||_2$ to both sides of (4) and using the triangle inequality yields

$$||f(x)||_2 \leq ||f(x_0)||_2 + ||\nabla f(x_0)||_2 \, ||\Delta x||_2 + ||T(x, x_0)||_2 \tag{20}$$

where $||\cdot||_2$ in $||\nabla f(x_0)||_2$ denotes the induced matrix norm. Combining (19) and (20) yields

$$||f(x)||_2 \leq ||f(x_0)||_2 + ||\nabla f(x_0)||_2 \, ||x||_2 + \frac{1}{2}\tau||\Delta x||_2^2 \tag{21}$$

with τ as defined in (14). By assumption (15) the r.h.s. of (21) is smaller than or equal to ρ for all $x \in B(\rho)$. Therefore $||f(x)||_2 \leq \rho$ for all $x \in B(\rho)$. Since for any $x_k \in B(\rho)$ this implies $||x_{k+1}||_2 = ||f(x_k)||_2 < \rho$, trajectories of (1) never leave $B(\rho)$ and therefore the dynamical system (1) is p.i. on $B(\rho)$.

Before proceeding to an example in section 4 we first discuss the complexity of calculating bounds $\underline{\lambda}_i$ and $\bar{\lambda}_i$ on the spectrum of Hessian matrices $\nabla^2 f_i(x)$.

3. HESSIAN MATRIX EIGENVALUE BOUNDS

Clearly, the computational complexity of the second order criterion in proposition 2 is higher than that of the first order criterion in proposition 1, since the calculation of eigenvalue bounds for the Hessian matrices, $\nabla^2 f_i(x)$, of each f_i, $i = 1, \ldots, n$ is required in the former, while eigenvalue bounds on only one Jacobian matrix, $\nabla f(x)$, is required for the latter. In this section we discuss the complexity of the calculation of eigenvalue bounds of Hessian matrices. To this end, we extend a result on the complexity of eigenvalue bounds for real Hessian matrices to interval Hessian matrices.

We briefly note that various other approaches to calculating eigenvalue bounds exist. Most notably, the calculation of tight eigenvalue bounds of an interval Hessian matrix, i.e. the calculation of the smallest possible upper bound and the largest possible lower bound, has been shown to be NP-hard (Blondel and Tsitsiklis, 2000). In fact tight bounds on the spectrum of an $n \times n$-dimensional symmetric interval matrix can be established by calculating the eigenvalues of 2^{n-1} real vertex matrices of the interval matrix (Hertz (1992); Rohn (1994)). Non-tight eigenvalue bounds can also be calculated with methods of lower complexity. For example, an interval variant of Gershgorin's circle method can be devised (Floudas, 1999). While this method in general does not yield tight bounds, it requires only $\mathcal{O}(n)$ arithmetic operations. Many other approaches with computational complexities between those of the interval Gershgorin and the vertex matrix approach exist. For an introduction, we refer the reader to Floudas (1999).

3.1 Preliminaries and notation

Before stating the necessary result on spectra of Hessians, we need to collect some preliminary results and introduce some notation. We carry out calculations with bounds on real numbers and borrow a compact notation from interval arithmetics (see Kearfott (1996), for example). Let $\underline{c}$ and $\bar{c} \geq \underline{c}$ denote real numbers. If a variable $c \in R$ is bounded according to $\underline{c} \leq c \leq \bar{c}$ this is denoted by $c \in [\underline{c}, \bar{c}]$ or $c \in [c]$, where $[c] := [\underline{c}, \bar{c}]$ for short. Note intervals represent real numbers if the lower and upper bounds are equal. If, for example $\underline{c} = \bar{c}$, the interval $[\underline{c}, \bar{c}]$ is identified with the real number $\underline{c} = \bar{c}$. Intervals of this type are called flat for short. Calculations with intervals can be carried out according to the following rules.

Fact 4. (Kearfott, 1996) Assume real numbers c and d are bounded according to $c \in [c]$ and $d \in [d]$ for some real intervals $[c] = [\underline{c}, \bar{c}]$ and $[d] = [\underline{d}, \bar{d}]$. Let $a \in R$ be an arbitrary real number. Then

$$\begin{aligned}
c + d &\in [c] + [d] := [\underline{c} + \underline{d}, \bar{c} + \bar{d}], \\
c - d &\in [c] - [d] := [\underline{c} - \bar{c}, \bar{c} - \underline{d}], \\
a\, c &\in a\,[c] := \begin{cases} [a\,\underline{c}, a\,\bar{c}] \text{ if } a \geq 0 \\ [a\,\bar{c}, a\,\underline{c}] \text{ if } a < 0 \end{cases} \\
c\, d &\in [c]\,[d] := [\min(\underline{c}\,\underline{d}, \underline{c}\,\bar{d}, \bar{c}\,\underline{d}, \bar{c}\,\bar{d}), \max(\underline{c}\,\underline{d}, \underline{c}\,\bar{d}, \bar{c}\,\underline{d}, \bar{c}\,\bar{d})]
\end{aligned} \tag{22}$$

If $0 \notin [d]$, the interval division $c/d \in [c]/[d]$, where $[c]/[d] := [\min(\underline{c}/\underline{d}, \underline{c}/\bar{d}, \bar{c}/\underline{d}, \bar{c}/\bar{d}), \max(\underline{c}/\underline{d}, \underline{c}/\bar{d}, \bar{c}/\underline{d}, \bar{c}/\bar{d})]$, is well-defined.

We state similar arithmetic rules for bounds on eigenvalues of symmetric matrices.

Lemma 5. Let $\mathcal{A}$ and $\mathcal{B}$ be sets of symmetric real matrices. Let $[\underline{c}, \bar{c}]$, $[\underline{\alpha}, \bar{\alpha}]$ and $[\underline{\beta}, \bar{\beta}]$ be intervals and assume

$$\lambda_A \in [\underline{\alpha}, \bar{\alpha}], \lambda_B \in [\underline{\beta}, \bar{\beta}]$$

for all eigenvalues λ_A and λ_B of all matrices $A \in \mathcal{A}$ and $B \in \mathcal{B}$, respectively. Then

$$\lambda \in [\underline{c}, \bar{c}]\,[\underline{\alpha}, \bar{\alpha}], \tag{23}$$

for all eigenvalues λ of all matrices from the set

$$\{C : C = c\,A \text{ for some } c \in [\underline{c}, \bar{c}] \text{ and some } A \in \mathcal{A}\}$$

and

$$\lambda \in [\underline{\alpha}, \bar{\alpha}] + [\underline{\beta}, \bar{\beta}] \tag{24}$$

for all eigenvalues λ of all matrices from the set

$$\{C : C = A + B \text{ for some } A \in \mathcal{A} \text{ and some } B \in \mathcal{B}\}.$$

The proof of lemma 5 is elementary and omitted for brevity.

Finally, we state the following lemma as an example for a Hessian matrix eigenvalue bound arithmetic rule. The particular composition of functions, $f(x) = h^3(x)$, treated in the lemma is chosen because a term of this form occurs in the example stated further below.

Lemma 6. Let $h : U \to R$ be a twice continuously differentiable function on an open $U \subseteq R^n$. Let $f(x) = h^3(x)$. Assume $S \subset U$ is compact and assume there exist real numbers $\underline{h}$, $\bar{h}$, $\underline{\lambda}_h$, $\bar{\lambda}_h$, and $\underline{\sigma}$, $\bar{\sigma}$ such that

$$\begin{aligned}
h(x) &\in [\underline{h}, \bar{h}] \text{ for all } x \in S, \\
\lambda &\in [\underline{\lambda}_h, \bar{\lambda}_h], \\
&\text{for all eigenvalues of } \nabla^2 h(x) \text{ for all } x \in S, \text{ and} \\
\sigma &\in [\underline{\sigma}, \bar{\sigma}] \\
&\text{for all eigenvalues } \sigma \text{ of } \nabla h(x)(\nabla h(x))^T \text{ for all } x \in S. \text{ Then} \\
\lambda_f &\in 3\,[\underline{h}, \bar{h}]\left(2\,[\underline{\sigma}, \bar{\sigma}] + [\underline{h}, \bar{h}]\,[\underline{\lambda}_h, \bar{\lambda}_h]\right)
\end{aligned} \tag{25, 26}$$

for all eigenvalues λ_f of $\nabla^2 f(x)$ for all $x \in S$.

Proof. Using the chain rule we get

$$\nabla^2 f(x) = 3\,h(x)\left(2\,\nabla h(x)(\nabla h(x))^T + h(x)\,\nabla^2 h(x)\right). \tag{27}$$

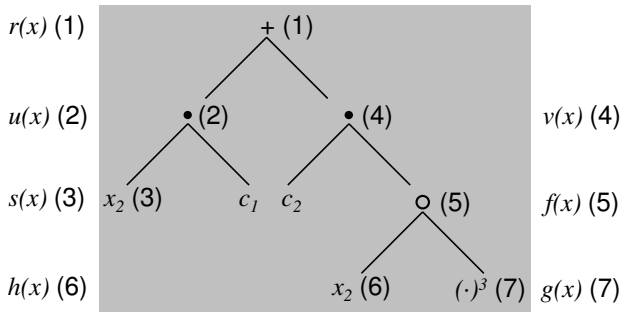

Fig. 1. Binary tree of the sample function $r(x) = c_1\, x_2 + c_2\, x_2^3$ where $c_1 = \frac{9}{10}$ and $c_2 = \frac{1}{10}$. The numbering of nodes is arbitrary. The symbol $\circ$ denotes the composition of two functions, i.e. $g(h(x)) = (g \circ h)(x)$. Function names to the left and right are not part of the binary tree but listed for ease of reference only.

By assumption the spectrum of $\nabla^2 h(x)$ and the function $h(x)$ are confined to $[\underline{\lambda}_h, \bar{\lambda}_h]$ and $[\underline{h}, \bar{h}]$ for all $x \in S$, respectively. Using rule (23) we infer that, for all $x \in S$, the eigenvalues of the matrix $(h(x)\, \nabla^2 h(x))$ are confined to

$$[\underline{h}, \bar{h}]\, [\underline{\lambda}_h, \bar{\lambda}_h]. \tag{28}$$

Applying rule (23) to the flat interval $[\underline{c}, \bar{c}] = [2, 2]$ and using assumption (25) we infer the eigenvalues of $2\, \nabla h(x) (\nabla h(x))^T$ to be confined to

$$2\, [\underline{\sigma}, \bar{\sigma}]. \tag{29}$$

Adding (28) and (29) according to rule (24) results in bounds

$$2\, [\underline{\sigma}, \bar{\sigma}] + [\underline{h}, \bar{h}]\, [\underline{\lambda}_h, \bar{\lambda}_h] \tag{30}$$

on the eigenvalues of the matrix defined by the term in the bracket on the r.h.s. of (27) for all $x \in S$. Using rule (22) we infer $3\, h(x) \in 3\, [\underline{h}, \bar{h}]$. Multiplying this result with (30) according to rule (23) gives the desired bounds (26).

3.2 Motivating example

The following example demonstrates how fact 4 and lemmata 5 and 6 can be used to calculated bounds on eigenvalues with a simple recursion. The purpose of this recursion is to split a function into simpler and simpler subfunctions until trivial subfunctions of the form $x \to x_i$ are obtained.

Example 7. Let $x = (x_1, x_2)^T \in R^2$, $r(x) = \frac{9}{10}\, x_2 + \frac{1}{10}\, x_2^3$, and $S = \left\{(x_1, x_2)^T \in R^2 : x_1 \in [-1/2, 1/2], x_2 \in [-1/2, 1/2]\right\}$. The function r can recursively be split into simpler functions. Specifically, $r(x) = u(x) + v(x)$ where $u(x) = \frac{9}{10}\, s(x)$, $s(x) = x_2$ and $v(x) = \frac{1}{10}\, g(h(x))$ where $h(x) = x_2$ and $g : R \to R$, $g(\xi) = \xi^3$. A natural representation of this recursive splitting is a binary tree, cf. figure 1.

Based on this binary tree, eigenvalue bounds can be determined without calculating a Hessian matrix of any but the trivial functions $s(x)$ and $h(x)$. In the numbering Nn used in what follows, n corresponds to the node number in figure 1.

N1 Assume the spectra of $\nabla^2 u(x)$ and $\nabla^2 v(x)$ are known to be confined to intervals $[\lambda_u] = [\underline{\lambda}_u, \bar{\lambda}_u]$ and $[\lambda_v] = [\underline{\lambda}_v, \bar{\lambda}_v]$ for all $x \in S$. Then for all eigenvalues λ_r of $\nabla^2 r(x)$ we can infer $\lambda_r \in [\lambda_u] + [\lambda_v]$ for all $x \in S$ according to rule (24). We have to descend into node 2 and node 4 to calculate $[\lambda_u]$ and $[\lambda_v]$, respectively.

N2 Assume the spectrum of $\nabla^2 s(x)$ is known to be confined to an interval $[\lambda_s]$ for all $x \in S$. Then for all eigenvalues λ_u of $\nabla^2 u(x)$ we can infer $\lambda_u \in \frac{9}{10}\, [\lambda_s]$ for all $x \in S$ using rule (23). We have to descend into node 3 to calculate $[\lambda_s]$.

N3 Since $s(x) = x_2$ implies

$$\nabla^2 s(x) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

all eigenvalues of $\nabla^2 s(x)$ are zero for all $x \in S$. Consequently, we have the trivial result $[\lambda_s] = [0, 0]$. Since the binary tree ends at node 3, no further descending is necessary.

N4 Assume the spectrum of $\nabla^2 f(x)$ is known to be confined to an interval $[\lambda_f]$ for all $x \in S$. Then the spectrum of $\nabla^2 v(x) = \frac{1}{10}\, \nabla^2 f(x)$ is confined to $[\lambda_v] = \frac{1}{10}\, [\lambda_f]$ according to rule (23). We have to descend into node 5 to calculate $[\lambda_f]$.

N5 Assume bounds $[\underline{h}, \bar{h}]$, $[\underline{\lambda}_h, \bar{\lambda}_h]$, and $[\underline{\sigma}, \bar{\sigma}]$ are known on the range of $h(x)$, the spectrum of $\nabla h(x)\, (\nabla h(x))^T$, and the spectrum of $\nabla^2 h(x)$ for all $x \in S$, respectively. Then bounds $[\lambda_f]$ on the spectrum of $\nabla^2 f(x) = \nabla^2 g(h(x))$ can be calculated according to (26) in lemma 6. We have to descend into node 6 to calculate $[h]$, $[\sigma]$, and $[\lambda_h]$.

N6 Nodes 3 and 6 are equal, therefore $[\lambda_h] = [0, 0]$. Since $x_2 \in [-1/2, 1/2]$ by assumption, $h(x) = x_2 \in [-1/2, 1/2]$. Since $\nabla h(x) = (0, 1)^T$,

$$\nabla h(x)\, (\nabla h(x))^T = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

This matrix has eigenvalues 0 and 1, therefore $[\sigma]$ can be set to $[\underline{\sigma}, \bar{\sigma}] = [0, 1]$ where $\underline{\sigma}$ and $\bar{\sigma}$ are as defined in lemma 6.

Using the results for the terminal nodes we can now proceed to calculate bounds on the spectra of the Hessian matrices of the functions defined at the other nodes of the tree. Using the result for node 6 and rule (N5) we infer

$$[\lambda_f] = 3 \left[-\frac{1}{2}, \frac{1}{2}\right] \left(2\, [0, 1] + \left[-\frac{1}{2}, \frac{1}{2}\right]\, [0, 0]\right) = [-3, 3].$$

Using this result and (N4) we find $[\lambda_v] = \frac{1}{10}\, [\lambda_f] = [-3/10, 3/10]$.

Using the result for node 3 and rule (N2) we find $[\lambda_u] = [0, 0]$. Using this result and the result for node 4, $[\lambda_v] = [-3/10, 3/10]$, we infer

$$[\lambda_r] = [-3/10, 3/10] \tag{31}$$

with rule (N1). This concludes the example.

Note this recursion by construction always ends at functions for which eigenvalue bounds can be found trivially.

3.3 Computational complexity

At first sight the calculation of the eigenvalue bounds for example 7 looks quite elaborate. A closer look reveals, however, that calculations of this type can easily be carried out automatically at a low computational cost.

In the remainder of this section we state a result on the computational complexity of an eigenvalue bound calculation.

Let $r : U \to R$ be an arbitrary function on an open $U \subset R^n$ that can be represented by a tree of binary and unary operations ω,

$$\omega \in S_\omega \{+, -, \times, /, \sin, \exp, \dots\}. \tag{32}$$

The set S_ω can be extended by any other binary or unary operation ω on real operands for which ω and its first and second derivative are well-defined, continuous, and can be implemented in a computer language. Other extensions are possible (Kearfott, 1996).

As a preparation we state a result on the complexity of the calculation of bounds on the spectrum of $\nabla^2 r(x_0)$ *at an arbitrary but fixed point* $x_0 \in U$. In what follows we carefully need to distinguish between two cases: (i) bounds on the spectrum of $\nabla^2 r(x)$ for all $x \in S$ in some hyperrectangle $S \subset U$, and (ii) bounds on the spectrum of $\nabla^2 r(x_0)$ at a particular point $x_0 \in U$. We denote the former bounds by $[\lambda_r]$ and the latter by $[\lambda_r]_{\text{point}}$. By $N(r)$ denote the number of operations $\omega \in S_\omega$ that are necessary to evaluate r at a point x_0. By $N([\lambda_r]_{\text{point}})$ denote the number of operations necessary to compute bounds on the spectrum of $\nabla^2 h(x_0)$ at a point x_0. Note $N(r)$ is determined by the number of nodes in the binary tree used for the computation of $r(x_0)$ but independent of the choice of the point x_0. Likewise, $N([\lambda_r]_{\text{point}})$ is determined by the number of nodes in the binary tree and the number of operations per node, but independent of x_0. The numbers $N(r)$ and $N([\lambda_r]_{\text{point}})$ are related as stated in the following lemma.

Lemma 8. (Mönnigmann, 2007) There exist positive integers α and β such that

$$\frac{N([\lambda_r]_{\text{point}})}{N(r)} = \alpha\, n + \beta, \tag{33}$$

for all twice-continuously differentiable $r(x)$ that can be represented by a tree of operations $\omega \in S_\omega$.

While the proof of lemma 8 is beyond the present paper, we give a sketch of it. At each node of the binary tree, one operation $\omega \in S_\omega$ needs to be carried out in order to evaluate the function at x_0. By $N(\nabla r(x_0), r(x_0))$ denote the number of operations $\omega \in S_\omega$ that are necessary to evaluate both the functions value $r(x)$ and the gradient $\nabla r(x)$ at x_0. From a standard result in automatic differentiation (Rall, 1981) we know

$$N(\nabla r(x_0), r(x_0)) = \mathcal{O}(n)\, N(r(x_0)).$$

On the other hand we know (Kearfott, 1996) that

$$N([\nabla r], [r]) = \mathcal{O}(1)\, N(\nabla r(x_0), r(x_0))$$

operations are necessary to compute interval enclosures of $\nabla r(x)$ and $r(x)$ on a hyperrectangle S in the domain of r. Using rules like (26) we can calculate bounds on the spectrum of the Hessian of each node. Since this calculation involves only interval arithmetics operations on real intervals but not on vectors or matrices of intervals (cf. (26)), the number of operations is independent of n. Therefore the number of operations necessary at each node increases by a factor $\mathcal{O}(1)$. In summary the complexity is

$$N([\lambda_r]_{\text{point}}) = \mathcal{O}(1)\, N([\nabla r], [r]) = \mathcal{O}(1)\, \mathcal{O}(n)\, N(r),$$

and this is the desired result (33).

Based on lemma 8 we can now state the desired result on the complexity of computing eigenvalue bounds for Hessian matrices on hyperrectangles.

Proposition 9. There exist positive integers $\tilde{\alpha}$ and $\tilde{\beta}$ such that

$$\frac{N([\lambda_r])}{N(r)} = \tilde{\alpha}\, n + \tilde{\beta} \tag{34}$$

for all twice continuously differentiable $r(x)$ that can be represented by a tree of operations $\omega \in S_\omega$.

Proof. The operations necessary in each node of the binary tree of r can be augmented by their corresponding interval arithmetics operations (Kearfott, 1996). At each node of the binary tree of r, the computational cost increases by a factor independent of n (Kearfott, 1996). Therefore $N([\lambda_r]) = \mathcal{O}(1)\, N([\lambda]_{\text{point}})$. Substituting into (33) we obtain the desired result.

Loosely speaking, proposition 9 tells us that we can compute bounds $[\lambda_r]$ on the spectrum of $\nabla^2 r(x)$ that hold for all $x \in S \subset R^n$ at a computational cost that is only linearly (in n) more expensive than the computation of $r(x)$ at a point $x \in S$. This result is surprising when compared to the complexity of computing an interval Hessian matrix, i.e. a matrix with elements

$$[\underline{H}_{ij}, \bar{H}_{ij}],\ i, j = 1, \dots, n$$

such that

$$\nabla^2 h(x)_{ij} \in [\underline{H}_{ij}, \bar{H}_{ij}] \tag{35}$$

for all $x \in S$ and all $i, j \in \{1, \dots, n\}$. The computational complexity of evaluating a Hessian matrix $\nabla^2 f_i(x)$ *at a point* x_0 is

$$\mathcal{O}(n^2)\, N(r)$$

if the same techniques from interval arithmetics and automatic differentiation are used as for the calculation of $[\lambda_r]$ (Rall, 1981). Since the use of interval arithmetics instead of floating point arithmetics increases this cost by a factor of order $\mathcal{O}(1)$ as in lemmata 8 and 9, the computational cost of calculating (35) is

$$\mathcal{O}(1)\, \mathcal{O}(n^2)\, N(r) = \mathcal{O}(n^2)\, N(r).$$

According to proposition 9, this is one order (in n) more expensive than the calculation of $[\lambda_r]$.

4. ILLUSTRATIVE EXAMPLE

In this section we show that systems exist such that p.i. can be shown for some domain with the second order criterion from proposition 2 but not with the first order criterion from proposition 1.

Consider system (1) where $n = 2$ and

$$\begin{aligned} f_1(x) &= \frac{9}{10}\, x_1 + \frac{1}{10}\, x_2^3, \\ f_2(x) &= \frac{9}{10}\, x_2 + \frac{1}{10}\, x_2^2 \end{aligned} \tag{36}$$

with fixpoint $x_0 = (0, 0)^T$. Since the Jacobian matrix of f,

$$A = \begin{pmatrix} \frac{9}{10} & \frac{3}{10} x_2^2 \\ 0 & \frac{9}{10} + \frac{1}{5} x_2 \end{pmatrix} \tag{37}$$

is an upper triangular matrix, its eigenvalues are equal to its diagonal elements, i.e. $\lambda_1 = 9/10$ and $\lambda_2 = 9/10 + 1/5\, x_2$. At the fixpoint $x = (0, 0)^T$ the eigenvalues evaluate

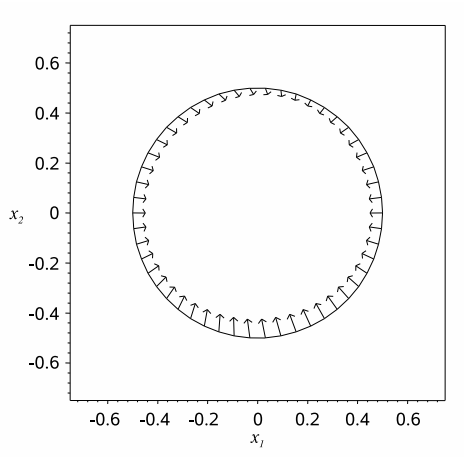

Fig. 2. One time step evolution of sample system (36) for some initial conditions on the boundary of $B(1/2)$.

to $\lambda_{1,2} = 9/10$. Since the eigenvalues lie strictly inside the unit circle, the fixpoint is locally stable.

Figure 2 shows single time steps $x_1 = f(x_0)$ for a number of initial points x_0 on the boundary of $B(1/2)$. This figure suggests that $B(1/2)$ is p.i. for the dynamical system (36). We show that the positive invariance on $B(1/2)$ can be established with the bounds on the eigenvalues of the Hessian matrices, but not with the bounds on the eigenvalues on the Jacobian matrix.

We begin with the linear order and therefore attempt to apply proposition 1. The eigenvalues of $\nabla f(x)^T \nabla f(x)$ evaluated at $x = (0,1)^T$ are $\lambda_1 \approx 0.83$ and $\lambda_2 \approx 1.19$ (calculation not shown). This implies $\sigma > 1$ for σ as defined in (7). For the point $(1/2, 0) \in B(1/2)$, for example, the l.h.s. of condition (8) evaluates to

$$\sigma \, ||\Delta x||_2 = \sigma \, \frac{1}{2} > \frac{1}{2}$$

Since $\rho = 1/2$, condition (8) is violated. Positive invariance of $B(1/2)$ can therefore not be inferred with the linear remainder condition stated in proposition 1.

In order to apply the second order criterion, we need bounds $\underline{\lambda}_i$, $\bar{\lambda}_i$ on the spectra of the Hessian matrices of f_1 and f_2. The function f_1 is equal to the function $r(x)$ treated in example 7. Since $B(1/2) \subset S$ for S as defined in example 7, the eigenvalues of $\nabla^2 f_1(x)$ are bounded by the interval given in (31), i.e.

$$[\underline{\lambda}_1, \bar{\lambda}_1] = [-3/10, 3/10].$$

Since f_2 as defined in (36) is quadratic, its Hessian matrix is independent of x and bounds on its eigenvalues can easily be obtained. Specifically,

$$H_2(x) = \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{5} \end{pmatrix}, \tag{38}$$

for all $x \in R$ with eigenvalues 0, 1/5. Therefore, $\underline{\lambda}_2 = 0$ and $\bar{\lambda}_2 = 1/5$. The quantity τ as defined in (14) is equal to $\tau = \sqrt{(3/10)^2 + (1/5)^2} = \sqrt{13}/10$. Finally, $||\nabla f(x_0)||_2 = \left(\lambda_{\max}(\nabla f(x_0)^T \nabla f(x_0))\right)^{1/2}$ is equal to 9/10 (calculation not shown). Having collected the necessary intermediate results, criterion (15) can be evaluated to give

$$\frac{9}{10}||\Delta x||_2 + \frac{1}{2}\frac{\sqrt{13}}{10}||\Delta x||_2^2 \leq 1/2 \tag{39}$$

for the example. Since $\sqrt{13}/10 < 0.361$ and since

$$\max_{x \in B(1/2)} ||\Delta x||_2 = 1/2 \text{ and } \max_{x \in B(1/2)} ||\Delta x||_2^2 = 1/4,$$

the l.h.s. of (39) is bounded above by $\frac{9}{20} + \frac{\sqrt{13}/4}{20}$ which is approximately equal to, but smaller than, 0.4951. Therefore the second order criterion (15) is fulfilled.

In summary we are able to show that $B(1/2)$ is p.i. for the example with the second order criterion and bounds on the eigenvalues of the Hessian matrices of the f_i, while we cannot establish positive invariance with the first order criterion and eigenvalue bounds on Jacobian of f.

5. CONCLUSION

We introduced two simple criteria for the positive invariance of balls around fixpoints of general nonlinear n-dimensional autonomous discrete time system. These criteria are based on well-known linear and quadratic remainders, respectively, for Taylor's theorem. We investigate the Taylor series expansion to second order, because the linear part of expansion to second order is globally stable if the original system is only locally stable. We demonstrated that systems exist for which the second order criterion can be used to infer positive invariance while the first order criterion fails.

Future work will have to address several extensions. The proposed approach is currently restricted to discrete time systems without inputs. Therefore an extension discrete time systems with inputs and to continuous time systems with and without inputs must be called for. Secondly, the conditions for positive invariance are sufficient but not necessary. Necessary conditions will become important in particular if an efficient computational procedure for testing for positive invariance is to be devised in the future.

The author thanks D. Kastsian for careful reading of the manuscript and one of the anonymous reviewers for pointing out a mistake in the first draft.

REFERENCES

F. Blanchini. Set invariance in control. *Automatica*, 35: 1747–1767, 1999.

V. D. Blondel and J. N. Tsitsiklis. A survey of computational complexity results in systems and control. *Automatica*, 36:1249–1274, 2000.

C. A. Floudas. *Deterministic Global Optimization: Theory, Methods, and Applications*. Kluwer, 1999.

D. Hertz. The extreme eigenvalues and stability of real symmetric interval matrices. *IEEE Trans. Autom. Control*, 37(4):532–535, 1992.

H. Heuser. *Lehrbuch der Analysis, Teil 2*. Teubner, 1991.

R. B. Kearfott. *Rigorous Global Search: Continuous Problems*. Kluwer Academic Publishers, 1996.

M. Mönnigmann. A new efficient approach to bounding spectra of Hessian matrices. Submitted, 2007.

L. B. Rall. *Automatic Differentiation: Techniques and Applications*. Springer, 1981.

J. Rohn. Positive definiteness and stability of interval matrices. *SIAM J. Matrix Anal. Appl.*, 15(1):175–184, 1994.