A new efficient eigenvalue bounding method for convexity detection with applications in global optimization and control

M. Mönnigmann
Department of Mechanical Engineering
Technische Universität Braunschweig, Braunschweig, Germany

September 17, 2008

We introduce a new method for the calculation of bounds on the eigenvalues of Hessian matrices of twice continuously differentiable functions. Eigenvalue bounds of Hessian matrices arise in a number of notoriously difficult tasks in computational chemical engineering. For example, Hessian matrix eigenvalue bounds are used in global nonlinear optimization, global convexity/concavity analysis in convex optimization, and global positive/negative invariance analysis in nonlinear control.

We stress that the improvements in computational complexity to be stated below are only possible, because the desired Hessian matrix eigenvalue bounds are calculated without ever calculating the Hessian matrix itself. To the author's knowledge the proposed method is the first Hessian-matrix-free approach to bounding Hessian matrix eigenvalues.

We start with a more precise problem statement in the next section, summarize the methodological advances in a subsequent section, and turn to applications in the last section.

Problem statement Two variants of the proposed method must be carefully distinguished from one another. Both variants apply to twice continuously differentiable functions $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$. Let $x_0 \in U$ be an arbitrary point and let $S = [\underline{x}_1, \bar{x}_1] \times \cdots \times [\underline{x}_n, \bar{x}_n]$ be an arbitrary closed hyperrectangle. With the first variant we calculate bounds on the eigenvalues of a Hessian matrix that has been evaluated *at a point*. More precisely, we determine $\underline{\lambda} \in \mathbb{R}$ and $\bar{\lambda} \in \mathbb{R}$ such that $\lambda \in [\underline{\lambda}, \bar{\lambda}]$ for all eigenvalues λ of $\nabla^2 f(x)$ evaluated at an arbitrary but fixed $x_0 \in U$. The second variant provides bounds that apply to all eigenvalues of the Hessian $\nabla^2 f(x)$ evaluated *on a hyperrectangle*. More precisely, we calculate bounds $\underline{\lambda} \in \mathbb{R}$ and $\bar{\lambda} \in \mathbb{R}$ such that $\lambda \in [\underline{\lambda}, \bar{\lambda}]$ for all eigenvalues λ of $\nabla^2 f(x)$ for all $x \in S$. For brevity we refer to these two cases as the real and the interval variant of the method, respectively. These names are chosen because the first variant applies to real matrices $\nabla^2 f(x) \in \mathbb{R}^{n \times n}$ while Hessian matrices on hyperrectangles are often approximated by interval Hessian matrices.

Advances and relation to existing methods Both the real and the interval variant of the new method are compared to established methods, specifically the method due to Gershgorin [3] and the method due to Hertz [4] and Rohn [6], respectively. Hertz and Rohn's method is known to provide *tight*, i.e. the best possible, bounds on the eigenvalues of interval matrices. We note that Hertz and Rohn's method is exponentially complex, however. This complexity is not surprising, since the problem of calculating tight eigenvalue bounds for interval matrices is known to be NP-hard [1].

In a formal proof the real variant of the method can be shown to be one order of magnitude computationally less complex than the method due to Gershgorin. More precisely, the proposed method provides bounds at a complexity

$$\mathcal{O}(n)\, N(f_i) \tag{1}$$

where $N(f_i)$ denotes the number of operations necessary to evaluate the function f_i at a point x_0, where $i \in \{1, \dots n\}$. In contrast, the calculation of the Hessian matrix and its eigenvalue bounds with Gershgorin's approach can be shown to require $\mathcal{O}(n^2)\, N(f_i)$ operations. Despite its lower complexity, nontrivial examples exist for which the proposed method results in bounds that are tighter than Gershgorin's bounds.

The results for the interval variant of the method are somewhat surprising and have to be summarized more carefully. First we note that the interval variant of the method can be shown to belong to the same polynomial complexity class (1) as its real counterpart. Furthermore, nontrivial examples of functions exist for which the new method provides bounds on the eigenvalues of the Hessian matrix on hyperrectangles that are *tighter* than the bounds obtained with Hertz and Rohn's method for the interval Hessian matrix. At first sight, this claim seems to be a contradiction to both the NP-hardness and the tightness of Hertz and Rohn's method. Closer inspection reveals that the restrictions that apply to methods for interval Hessians do *not* apply here, because neither real nor interval Hessian matrices are ever calculated in the proposed method.

Due to its polynomial complexity the interval variant of the new approach can be expected to be of use in cases where the exponential complexity of Hertz and Rohn's method is prohibitive.

Applications Bounds on eigenvalues of Hessian matrices can be used in a variety of applications. Eigenvalues of real Hessian matrices can be used to check sufficient optimality conditions in nonlinear programming, for example. Bounds on eigenvalues of Hessians on hyperrectangles can be used in deterministic approaches to global nonlinear optimization to create convex underestimators. A famous example for such an approach is the αBB method developed by Floudas and coworkers [2].

In the present contribution we use an application from stability and control theory to illustrate the use of the new method for the calculation of eigenvalue bounds. In stability and control theory, positive definiteness, or equivalently convexity, of quadratic forms appears in various criteria that are based on Lyapunov functions (see e.g. [5]).

Here we focus on the use of Hessian matrices on hyperrectangles for the search of positive invariant sets of dynamical systems. Essentially, a set in the domain of a nonlinear dynamical system is called positive invariant if the system never leaves it once it entered it. We state two simple criteria for positive invariance, which are based on the first and second order Taylor approximation of the nonlinear dynamical system. We show that examples exist where positive invariance can be established with the second order but not the first order criterion. In order to efficiently evaluate the second order criterion, bounds on eigenvalues of Hessian matrices are calculated with the new method proposed here.

References

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

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

[3] S. Gerschgorin. Über die Abgrenzung der Eigenwerte einer Matrix. *Izv. Akad. Nauk SSSR, Ser. fizmat.*, 6:749–754, 1931.

[4] D. Hertz. The extreme eigenvalues and stability of real symmetric interval matrices. *IEEE Transactions on Automatic Control*, 37(4):532–535, 1992.

[5] D. Hinrichsen and A. Pritchard. *Mathematical Systems Theory I*. Springer, 2005.

[6] J. Rohn. Positive definiteness and stability of interval matrices. *SIAM Journal on Matrix Analysis and Applications*, 15(1):175–184, 1994.