Struct Multidisc Optim (2014) 50:1197–1199
DOI 10.1007/s00158-014-1204-x

DISCUSSION

On the reduced Hessian of the compliance

Anton Evgrafov

Received: 24 November 2014 / Revised: 25 November 2014 / Accepted: 27 November 2014 / Published online: 21 December 2014

Abstract We describe simple numerical tests, which could have been used for verifying the derivation of the second order sensitivity analysis in a recent educational article "An efficient 3D topology optimization code written in Matlab" by Liu and Tovar (Struct Multidiscip Optim, 2014. doi:10.1007/s00158-014-1107-x). We also discuss the second order sensitivity analysis for the problem considered in the cited paper.

Keywords Second order sensitivity analysis · Hessian of the compliance

1 Introduction

The main purpose of this note is to discuss the second order sensitivity analysis presented in a recent educational article (Liu and Tovar 2014). Throughout this note we will utilize the notation employed in the cited article.

2 Discussion

In (Liu and Tovar 2014, equation (40) and Appendix B) it is claimed that the Hessian of the compliance with respect to the filtered density variables is a diagonal matrix, with non-negative diagonal elements given by

$$\frac{\partial^2 c}{\partial \tilde{x}_i^2} = 2\frac{\left[p\tilde{x}_i^{p-1}(E_0 - E_{\min})\right]^2}{E_{\min} + \tilde{x}_i^p(E_0 - E_{\min})}\mathbf{u}_i^{\mathrm{T}}\mathbf{k}_i^0\mathbf{u}_i, \qquad i = 1, \ldots, n. \tag{1}$$

If this were indeed the case, the compliance c would have been a convex and additively separable function of $\tilde{\mathbf{x}}$. As a result, for all vectors $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ and every constant $0 \le \lambda \le 1$ the following must hold:

$$c(\lambda\tilde{\mathbf{x}} + (1-\lambda)\tilde{\mathbf{y}}) \le \lambda c(\tilde{\mathbf{x}}) + (1-\lambda)c(\tilde{\mathbf{y}}), \qquad \text{and} \tag{2}$$

$$c(\tilde{\mathbf{x}}) = c(\tilde{x}_1, \ldots, \tilde{x}_n) = \sum_{i=1}^{n} c_i(\tilde{x}_i), \tag{3}$$

for some functions $c_i : \mathbb{R} \to \mathbb{R}$, $i = 1, \ldots n$, of one variable only. The immediate consequence of (3) is the so-called rectangle condition, that is, $\forall\tilde{\mathbf{x}} \in \mathbb{R}^n$, $\delta_i, \delta_j \in \mathbb{R}$, $i \ne j$ it holds that

$$c(\tilde{\mathbf{x}}) + c(\tilde{\mathbf{x}} + \delta_i\mathbf{e}_i + \delta_j\mathbf{e}_j) = c(\tilde{\mathbf{x}} + \delta_i\mathbf{e}_i) + c(\tilde{\mathbf{x}} + \delta_j\mathbf{e}_j), \tag{4}$$

where $\mathbf{e}_i, \mathbf{e}_j$ are the i^{th} and j^{th} canonical basis vectors in $\mathbb{R}^n$.

Conditions (2) and (4) can be easily tested numerically with the help of the code provided in (Liu and Tovar 2014). Utilizing the default example in the code (cantilever beam) with `nelx=60`, `nely=20`, `nelz=4`, `volfrac=0.3`, `penal=3` we test (4) on $\tilde{\mathbf{x}} = (0.3, \ldots, 0.3)^{\mathrm{T}}$, $i = 1$, $j = 21$ (this choice corresponds to the elements with indices (1, 1, 1) and (2, 1, 1) respectively), and $\delta = 0.7$. We find that $c(\tilde{\mathbf{x}}) \approx 1.4177 \cdot 10^4$, $c(\tilde{\mathbf{x}} + \delta\mathbf{e}_i) \approx 1.4100 \cdot 10^4$, $c(\tilde{\mathbf{x}} + \delta\mathbf{e}_j) \approx 1.4119 \cdot 10^4$, $c(\tilde{\mathbf{x}} + \delta\mathbf{e}_i + \delta\mathbf{e}_j) \approx 1.3991 \cdot 10^4$. Thus the residual of (4) at this point is ≈ 50.1, or in relative

Electronic supplementary material The online version of this article (doi: 10.1007/s00158-014-1204-x) contains supplementary material, which is available to authorized users.

A. Evgrafov (✉)
Department of Mathematical Sciences, Norwegian University of Science and Technology, N–7491 Trondheim, Norway
e-mail: anton.evgrafov@math.ntnu.no

terms $\approx 3.5 \cdot 10^{-3}$, far away from the accuracy expected from IEEE double precision floating point computations.

Similarly, putting $\lambda = 1/2$, $\tilde{\mathbf{x}}$ ($\tilde{\mathbf{y}}$) to be vectors of ones except at the elements (1, 1, 1) and (1, 1, 4) (respectively, (1, 1, 2) and (1, 1, 3)) where we put the value 0.01 we compute $c((\tilde{\mathbf{x}} + \tilde{\mathbf{y}})/2) \approx 387.99$, while $c(\tilde{\mathbf{x}}) \approx 385.45$ and $c(\tilde{\mathbf{y}}) \approx 385.25$. This results in an absolute violation of the inequality (2) by ≈ 2.65, or a relative violation of $\approx 6.8 \cdot 10^{-3}$.

Thus both of these simple numerical tests contest the formulae derived in (Liu and Tovar 2014, Appendix B), which contain several mistakes. First of all, the middle term in (Liu and Tovar 2014, equation (B2)), which is obtained by differentiating (Liu and Tovar 2014, equation (B1)) is guaranteed to be zero only when $i \neq j$ *or* $p = 1$. Otherwise it contributes non-positively to the diagonal of the reduced Hessian, and thereby to non-convexity of c, with the value

$$-\mathbf{u}_i^{\mathrm{T}} \frac{\partial^2 \mathbf{k}_i}{\partial \tilde{x}_i^2} \mathbf{u}_i = -p(p-1)\tilde{x}_i^{p-2}(E_0 - E_{\min})\mathbf{u}_i^{\mathrm{T}} \mathbf{k}_i^0 \mathbf{u}_i. \qquad (5)$$

Secondly, the equation

$$\mathbf{k}_i \mathbf{u}_i = \mathbf{f}_i, \qquad i = 1, \ldots, n, \qquad (6)$$

which forms the basis for equation (40) and other derivations in Appendix B of (Liu and Tovar 2014), is not implied by (Liu and Tovar 2014, equation (16)). If it were, the structural elements would have been completely decoupled from one another. Again, one may easily use the supplied code to verify this condition. Evaluating the second norm of the residual of (6) for all elements of the cantilever beam as described above at $\tilde{\mathbf{x}} = (0.3, \ldots, 0.3)^{\mathrm{T}}$ we obtain a median value of ≈ 0.18 (or even ≈ 2.6 when the residual is measured relative to the 2nd norm of $\mathbf{k}_i$).

Perhaps it is even more important to emphasize that even if (6) were true, the elemental matrix $\mathbf{k}_i$ *on the left hand side is singular*, which may again be checked numerically with the supplied code: `min(abs(eig(KE)))/norm(KE,2)` $\sim 10^{-17}$. Even though one may differentiate these equations, as the differentiability of $\mathbf{U}$ (hence also $\mathbf{u}_i$) is guaranteed by the inverse function theorem applied to the state equation (16) in (Liu and Tovar 2014), the equations (B3) through (B4) in (Liu and Tovar 2014) involving the inverse of $\mathbf{k}_i$ are meaningless in their present form.

At last but not the least, one could have used finite differences in connection with the Matlab code supplied with (Liu and Tovar 2014) to assess the correctness of the derived second order sensitivities.

3 Derivation of the Hessian of the compliance

Differentiating the equilibrium equation (Liu and Tovar 2014, equation (16)) twice, one obtains the expressions

$$\begin{aligned}
\frac{\partial^2 \mathbf{U}}{\partial \tilde{x}_i \partial \tilde{x}_j} &= -\mathbf{K}^{-1}\left[\frac{\partial^2 \mathbf{K}}{\partial \tilde{x}_i \partial \tilde{x}_j}\mathbf{U} + \frac{\partial \mathbf{K}}{\partial \tilde{x}_i}\frac{\partial \mathbf{U}}{\partial \tilde{x}_j} + \frac{\partial \mathbf{K}}{\partial \tilde{x}_j}\frac{\partial \mathbf{U}}{\partial \tilde{x}_i}\right], \text{ and} \\
\frac{\partial^2 c}{\partial \tilde{x}_i \partial \tilde{x}_j} &= \mathbf{F}^{\mathrm{T}} \frac{\partial^2 \mathbf{U}}{\partial \tilde{x}_i \partial \tilde{x}_j} \\
&= -\mathbf{U}^{\mathrm{T}}\left[\frac{\partial^2 \mathbf{K}}{\partial \tilde{x}_i \partial \tilde{x}_j}\mathbf{U} + \frac{\partial \mathbf{K}}{\partial \tilde{x}_i}\frac{\partial \mathbf{U}}{\partial \tilde{x}_j} + \frac{\partial \mathbf{K}}{\partial \tilde{x}_j}\frac{\partial \mathbf{U}}{\partial \tilde{x}_i}\right]. \qquad (7)
\end{aligned}$$

The first term on the right hand side of (7) contributes only to the diagonal of the Hessian matrix. It reduces to (5) when $i = j$ and therefore may be efficiently evaluated at the elemental level. Similarly, the vectors $[\partial \mathbf{K}/\partial \tilde{x}_i]\mathbf{U}$ are only non-zero at the degrees of freedom, corresponding to the element i and are equal to the elemental vector $[\partial \mathbf{k}_i/\partial \tilde{x}_i]\mathbf{u}_i = p\tilde{x}_i^{p-1}(E_0 - E_{\min})\mathbf{k}_i^0 \mathbf{u}_i$. Unfortunately, the vector $\partial \mathbf{U}/\partial \tilde{x}_j$ given by (Liu and Tovar 2014, equation (25)) is dense and its evaluation requires solving an additional linear system, for every $j = 1, \ldots, n$. The conclusion is that the reduced Hessian of the compliance is a large (in fact, extremely large in 3D) dense matrix and its computation is hardly a viable operation owing to the enormous memory requirements.

Instead, one can utilize a Krylov subspace solver for finding typical SQP-search directions, see for example (Liu and Tovar 2014, equation (38)). This approach allows one to mitigate the unrealistic memory requirements associated with storing $\nabla^2_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}} c$ by supplying a matrix-vector multiplications routine, which requires only sparse operations. Indeed, for a given $\mathbf{y} \in \mathbb{R}^n$ let us compute one component of the Hessian-vector product, $z_i = \sum_{j=1}^{n} \left[\partial^2 c/\partial \tilde{x}_i \partial \tilde{x}_j\right] y_j$. A simple rearrangement of (7) results in the following formula, whose correctness has been numerically verified using finite-differences:

$$z_i = -y_i \mathbf{U}^{\mathrm{T}} \frac{\partial^2 \mathbf{K}}{\partial \tilde{x}_i^2}\mathbf{U} + 2\left[\frac{\partial \mathbf{K}}{\partial \tilde{x}_i}\mathbf{U}\right]^{\mathrm{T}} \left\{\mathbf{K}^{-1}\left[\sum_{j=1}^{n} y_j \frac{\partial \mathbf{K}}{\partial \tilde{x}_j}\mathbf{U}\right]\right\}.$$

Thus only one additional system (the one in the curly brackets) needs to be solved per Hessian-vector multiplication. Furthermore, both expressions in the square brackets may be efficiently evaluated at the elemental level.

4 Alternatives to the "nested" approach

In conclusion we would like to mention that instead of treating the state variables $\mathbf{U}$ as an implicit function of the density variables $\tilde{\mathbf{x}}$ one may view all of them as "equal-rights" optimization variables. Such a view results in so called simultaneous analysis and design, one-step, or full space approaches, see for example (Hoppe et al. 2002; Schulz 2004; Othmer 2008). Within this approach the sparsity of the Hessian is inherited from the sparsity of the utilized PDE discretization.

Yet another alternative is to reverse the roles of $\mathbf{U}$ and $\tilde{\mathbf{x}}$, which for certain topology optimization problems results in very efficient algorithms based on the exact sparse Hessians (Carlsson et al. 2009; Evgrafov 2014).

References

Liu K, Tovar A (2014) An efficient 3D topology optimization code written in Matlab. Struct Multidiscip Optim. doi:10.1007/s00158-014-1107-x

Hoppe R, Petrova S, Schulz V (2002) A primal-dual Newton-type interior-point method for topology optimization. J Optimiz Theory App 114(3):545–571

Schulz V (2004) Simultaneous solution approaches for large optimization problems. J Comput Appl Math:629–641

Othmer C (2008) A continuous adjoint formulation for the computation of topological and surface sensitivities of ducted flows. Int J Numer Meth Fl 58:861–877

Carlsson J, Sandberg M, Szepessy A (2009) Symplectic Pontryagin approximations for optimal design. ESAIM-Math Model Num 43:3–32

Evgrafov A (2014) State space Newton method for topology optimization. Comput Method Appl M 278:272–290