

Jan Chleboun

An application of the averaged gradient technique

In: Jan Chleboun and Petr Příkryl and Karel Segeth and Tomáš Vejchodský (eds.): Programs and Algorithms of Numerical Mathematics, Proceedings of Seminar. Dolní Maxov, June 1-6, 2008. Institute of Mathematics AS CR, Prague, 2008. pp. 65–70.

Persistent URL: <http://dml.cz/dmlcz/702857>

**Terms of use:**

© Institute of Mathematics AS CR, 2008

Institute of Mathematics of the Czech Academy of Sciences provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This document has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library*  
<http://dml.cz>

# AN APPLICATION OF THE AVERAGED GRADIENT TECHNIQUE\*

Jan Chleboun

Dedicated to Ivan Hlaváček on the occasion of his 75th birthday.

## 1. Introduction

Gradient averaging (also known as gradient recovery (GR)) is a technique for improving the accuracy of an approximate gradient obtained via a numerical method. Sensitivity analysis deals with analyzing the response of a function (or a functional) to a small perturbation of its input values. In this contribution, we limit ourselves to gradients originating from finite element solutions of boundary value problems (BVPs) and to criterion-functionals that evaluate these solutions. Our goal is to show that the use of a gradient recovery technique in sensitivity analysis formulae can result in a better assessment of the quality of approximate minimizers of criterion-functionals that appear in parameter identification problems or the worst scenario method, for example.

Let us finish this short introductory part with a few words to honor Ivan Hlaváček from the Institute of Mathematics of the Academy of Sciences of the Czech Republic who recently celebrated his 75th birthday. He has pioneered a mathematically rigorous analysis of the worst scenario problems since the mid-nineties and also contributed to the family of gradient averaging techniques. The following pages thus pay tribute to his scientific achievements.

## 2. Averaged gradient

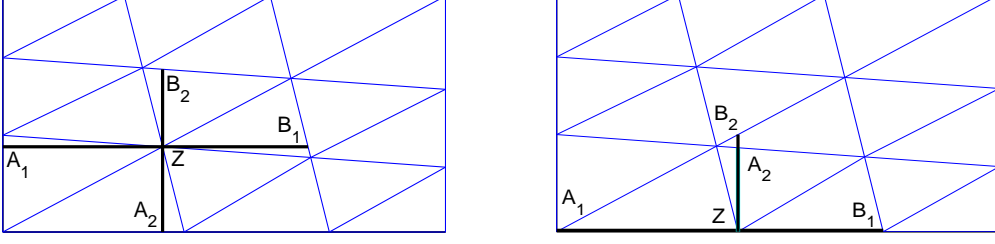
The idea to improve the accuracy of an approximate gradient calculated by the finite element method is more than two decades old and has materialized in numerous applications. Take, for example, Zienkiewicz-Zhu error estimators stemming from [8]. The contribution of Czech mathematicians is not negligible, see [2, 3, 4, 5, 6], for instance.

Although various recovery techniques have been designed, let us confine ourselves to a simple averaging method proposed and analyzed in [3].

Let  $\Omega \subset \mathbb{R}^2$  be a bounded domain with a polyhedral Lipschitz boundary. Let  $\mathcal{F} = \{T_h\}_{h \rightarrow 0}$  be a family of triangulations of  $\bar{\Omega}$ , where  $h$  is the maximum diameter

---

\*This research was supported by the Ministry of Education, Youth, and Sports of the Czech Republic through contract MSM 6840770003. The author's participation in the PANM 14 Conference was made possible through grant No. IAA100190803 from the Grant Agency of AS CR.



**Fig. 1:** Auxiliary points: for an inner node (left), for a boundary node (right).

of all elements  $K \in T_h$ . Let  $V_h = \{v_h \in C(\bar{\Omega}) : v_h|_K \in P_1(K) \ \forall K \in T_h\}$ , where  $P_1(K)$  stands for the space of linear polynomials on  $K$ .

For a mesh node  $Z$ , we draw two lines parallel to the axes (see Figure 1), find their intersections with the edges of those triangles that share the vertex  $Z$ , and label these intersection points  $A_1$ ,  $A_2$ ,  $B_1$ , and  $B_2$  as in Figure 1. We then set

$$a_i = (A_i - Z)_i, \quad b_i = (B_i - Z)_i, \quad i = 1, 2.$$

For  $v_h \in V_h$ , the components of the weighted averaged gradient  $G_h v_h \in V_h$  at  $Z$  are defined as follows

$$(G_h v_h(Z))_i = \alpha_i v_h(A_i) - (\alpha_i + \beta_i) v_h(Z) + \beta_i v_h(B_i), \quad (1)$$

where  $i = 1, 2$  and  $\alpha_i = b_i / (a_i(b_i - a_i))$ ,  $\beta_i = a_i / (b_i(a_i - b_i))$ .

If  $v \in C(\bar{\Omega})$ , then (1) can also be applied (with  $v_h$  replaced by  $v$ ) to define  $G_h v \in V_h$ , a continuous piece-wise linear approximation of  $\nabla v$ .

We refer to [3] for details and a generalization to  $\mathbb{R}^3$  as well as for situations that are not covered by Figure 1.

The features of  $\mathcal{F}$  are substantial for the order of accuracy of  $G_h v$ . Let us recall that  $\mathcal{F}$  is called a *strongly regular* family of triangulations if

$$\exists \varkappa > 0 \quad \forall T_h \in \mathcal{F} \quad \forall K \in T_h \quad \varkappa h \leq \varrho_K,$$

where  $\varrho_K$  is the radius of the largest ball inscribed in  $K$ .

It is known, see [3, Theorem 3.8], that if  $q \in (1, \infty)$  and  $\mathcal{F}$  is strongly regular, then a constant  $C > 0$  exists such that for any  $v$  belonging to the Sobolev space  $W_q^3(\Omega)$ , the following estimate holds (note the order  $h^2$ )

$$\|\nabla v - G_h v\|_{0,q,\Omega} \leq Ch^2 |v|_{3,q,\Omega} \quad \forall T_h \in \mathcal{F}. \quad (2)$$

Let us focus on the recovered gradient of a finite element (FE) solution.

We consider  $u \in V_0 = H_0^1(\Omega)$  ( $W_2^1(\Omega)$ -functions with zero trace), a unique weak solution to the following  $V_0$ -elliptic BVP

$$-\operatorname{div}(\tilde{\lambda} \nabla u) = f \quad \text{in } \Omega, \quad (3)$$

$$u = 0 \quad \text{on } \partial\Omega, \quad (4)$$

where  $\tilde{\lambda} \in W_{2+\varepsilon}^2(\Omega)$  for some  $\varepsilon > 0$  and  $f \in W_q^1(\Omega)$ . Let us remark that  $\tilde{\lambda}$  can be a symmetric matrix of functions, see [3]. It is assumed that  $u \in W_q^3(\Omega)$ , where  $q > 2$ .

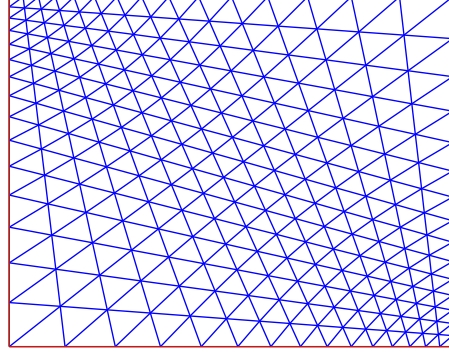
Next,  $u_h \in V_h^0 = V_0 \cap V_h$ , the piece-wise linear FE approximation of  $u$ , is defined through

$$\int_{\Omega} \tilde{\lambda}(x) \nabla u_h(x) \cdot \nabla v_h(x) \, dx = \int_{\Omega} f(x) v_h(x) \, dx \quad \forall v_h \in V_h^0. \quad (5)$$

To obtain a result similar to (2), we have to resort to more stringent assumptions about the meshes.

Let us consider  $\mathcal{F}' \subset \mathcal{F}$ , a special class of smoothly distorted uniform meshes, see [7, 3] for details and Figure 2 for an illustration of such a mesh.

For these meshes and for a fixed subdomain  $\Omega_0 \subset \subset \Omega$ , an analogue to (2) holds, see [3, page 22],



**Fig. 2:** *Distorted uniform mesh.*

$$\|\nabla u - G_h u_h\|_{0,2,\Omega_0} \leq C(u) h^2. \quad (6)$$

### 3. A parameter identification application

In this section, we will introduce a parameter identification problem.

For simplicity, let  $\tilde{\lambda}$  be controlled by six parameters forming a vector  $\lambda$ , that is,  $\lambda = (\lambda_1, \dots, \lambda_6)$  and

$$\tilde{\lambda}(x) = \lambda_1 + \lambda_2 x_1 + \lambda_3 x_2 + \lambda_4 x_1^2 + \lambda_5 x_1 x_2 + \lambda_6 x_2^2, \quad x = (x_1, x_2) \in \Omega.$$

We define the criterion-functional (and its approximation) that appears in parameter identification problems:

$$\Psi(\lambda) = \int_{\Omega} |\nabla u(\lambda) - \nabla w|^2 \, dx \quad \text{and} \quad \Psi_h(\lambda) = \int_{\Omega} |\nabla u_h(\lambda) - \nabla w|^2 \, dx,$$

where  $\Psi$  evaluates  $u(\lambda)$ , the  $\lambda$ -dependent weak solution to (3)–(4), and  $\Psi_h$  evaluates  $u_h(\lambda)$  determined by (5). In both cases,  $w$  is a given function.

Let  $\lambda \in \mathcal{U}_{\text{ad}} \subset \mathbb{R}^6$ , where  $\mathcal{U}_{\text{ad}}$  is a set of admissible parameters. The specification of  $\mathcal{U}_{\text{ad}}$  is not necessary for the purposes of this paper; roughly speaking,  $\mathcal{U}_{\text{ad}}$  is a compact subset of  $\mathbb{R}^6$  such that the BVP is uniformly  $V_0$ -elliptic with respect to  $\lambda \in \mathcal{U}_{\text{ad}}$ .

We formulate a parameter identification problem and its approximation: Find  $\lambda_0 \in \mathcal{U}_{\text{ad}}$  and  $\lambda_0^h \in \mathcal{U}_{\text{ad}}$  such that  $\lambda_0 = \arg \min_{\lambda \in \mathcal{U}_{\text{ad}}} \Psi(\lambda)$  and  $\lambda_0^h = \arg \min_{\lambda \in \mathcal{U}_{\text{ad}}} \Psi_h(\lambda)$ .

An elementary sensitivity analysis (see [1]) results in

$$\frac{\partial \Psi_h}{\partial \lambda_i}(\lambda) = - \int_{\Omega} \frac{\partial \tilde{\lambda}}{\partial \lambda_i} \nabla u_h \cdot \nabla z_h \, dx, \quad i = 1, 2, \dots, 6, \quad (7)$$

where  $z_h \in V_h^0$  is the solution to the adjoint equation

$$\int_{\Omega} \tilde{\lambda} \nabla z_h \cdot \nabla v_h \, dx = \int_{\Omega} 2(\nabla u_h - \nabla w) \cdot \nabla v_h \, dx \quad \forall v_h \in V_h^0(\Omega). \quad (8)$$

Computational experiments show that a gradient minimization procedure based on the derivative (7) is quite efficient in the search for  $\lambda_0^h$ . Nevertheless, the degree of accuracy of  $\lambda_0^h$  and  $\Psi_h(\lambda_0^h)$  remains unknown.

To obtain at least an indicator of the (in)accuracy of the minimization results, we will apply the above-mentioned gradient recovery technique.

Let us define both a new functional

$$\Psi_h^G(\lambda) = \int_{\Omega} |G_h u_h(\lambda) - \nabla w|^2 \, dx, \quad (9)$$

where  $u_h = u_h(\lambda)$  solves (5), and a new equation determining  $z_h^G \in V_h^0$  through

$$\int_{\Omega} \tilde{\lambda} \nabla z_h^G \cdot \nabla v_h \, dx = \int_{\Omega} 2(G_h u_h - \nabla w) \cdot \nabla v_h \, dx \quad \forall v_h \in V_h^0. \quad (10)$$

Strictly speaking, (10) is not the exact adjoint equation to (9) and (5) because the right-hand side of (10) is not the exact derivative of  $\Psi_h^G$  with respect to  $u_h$ ; see [1] for the derivation of adjoint equations. As a consequence,  $z_h^G$  does not yield the exact derivative of  $\Psi_h^G$ . However, the approximation, i.e.,

$$\frac{\partial \Psi_h^G}{\partial \lambda_i}(\lambda) \approx - \int_{\Omega} \frac{\partial \tilde{\lambda}}{\partial \lambda_i} \nabla u_h(\lambda) \cdot \nabla z_h^G(\lambda) \, dx, \quad i = 1, 2, \dots, 6,$$

is sufficiently accurate to be used in solving the following minimization problem: Find

$$\lambda_0^{h,G} = \arg \min_{\lambda \in \mathcal{U}_{\text{ad}}} \Psi_h^G(\lambda). \quad (11)$$

We end up with two approximate minimum points, that is,  $\lambda_0^h$  and  $\lambda_0^{h,G}$ , with two respective approximate state solutions  $u_h(\lambda_0^h)$  and  $u_h(\lambda_0^{h,G})$ , and three approximate criterion-functional values, namely  $\Psi_h(\lambda_0^h)$ ,  $\Psi_h^G(\lambda_0^h)$ , and  $\Psi_h^G(\lambda_0^{h,G})$ ; the fourth value,  $\Psi_h(\lambda_0^{h,G})$ , is not relevant for our purposes.

The distances

$$\sigma_{\lambda} = \left| \lambda_0^h - \lambda_0^{h,G} \right|, \quad \sigma_{\Psi} = \left| \Psi_h(\lambda_0^h) - \Psi_h^G(\lambda_0^{h,G}) \right|, \quad \text{and} \quad \sigma_{\Psi^G} = \left| \Psi_h^G(\lambda_0^h) - \Psi_h^G(\lambda_0^{h,G}) \right|$$

can indicate an inaccuracy in the approximation of the exact solution pair  $\lambda_0$  and  $\Psi(\lambda_0)$ . For  $h \rightarrow 0_+$ , it should be  $\sigma_{\lambda}, \sigma_{\Psi}, \sigma_{\Psi^G} \rightarrow 0$ .

Let us remark that there is no guarantee that  $\lambda_0^{h,G}$  approximates  $\lambda_0$  better than  $\lambda_0^h$  does.

*Example*

The problem is defined by  $\Omega = (-3, 3) \times (-3, 3)$ ,  $w = \exp(-x_1^2 - x_2^2)$ , and  $\Psi(\lambda) = 1000\|\nabla(u - w)\|_{0,2,\Omega}^2$ . The right-hand side function  $f$ , see (3), is determined by  $w$  and a predefined polynomial  $\hat{\lambda}$ . Although  $w$  does not comply with the homogeneous boundary condition, the difference is rather small and, for  $\hat{\lambda}$ , the state solution  $u$  is close to  $w$ . A distorted mesh is formed by 800 triangles ( $h = 0.3$ ).

The minimization process starts at  $\lambda_s = (2, 2, 2, 2, 2, 2)$ ; we obtain  $\Psi_h(\lambda_s) = 1641$  (GR not used) and  $\Psi_h^G(\lambda_s) = 1693$  (GR is used).

The calculated results are as follows

$$\begin{aligned} \Psi_h(\lambda_0^h) &= 93.9 && \text{GR is not used in the minimization,} \\ \Psi_h^G(\lambda_0^h) &= 26.5 && \text{GR is applied just to } u_h(\lambda_0^h), \\ \Psi_h^G(\lambda_0^{h,G}) &= 3.9 && \text{GR is used in the minimization,} \\ \sigma_\lambda &= 0.21, && \sigma_\Psi = 90.0, && \sigma_\Psi^G = 22.6. \end{aligned}$$

The minimum value  $\Psi_h(\lambda_0^h) = 93.9$  is the result of the minimization procedure where  $\nabla u_h(\lambda)$  is piece-wise constant. There is no indication whether or not  $\Psi_h(\lambda_0^h)$  is close to  $\Psi(\lambda_0)$ . If the averaged gradient of  $u_h(\lambda_0^h)$  is evaluated by the criterion-functional, we obtain  $\Psi_h^G(\lambda_0^h) = 26.5$ , which shows that  $G_h u_h$  is a better approximation of  $\nabla w$ . This can be anticipated because  $\nabla w$  is most significant in a subdomain  $\Omega_0 \subset\subset \Omega$  and (6) holds. The minimization based on (9) leads to even lower minimum  $\Psi_h^G(\lambda_0^{h,G}) = 3.9$ , which says that the averaged gradient of the state solution  $u_h(\lambda_0^{h,G})$  is so far the best approximation to  $\nabla w$ . Nevertheless,  $\nabla u_h(\lambda_0^{h,G})$  (piece-wise constant) does not outweigh  $\nabla u_h(\lambda_0^h)$  because the latter results in the criterion minimum value without GR. As indicated by  $\sigma_\lambda = 0.21$ , the difference between  $\lambda_0^h$  and  $\lambda_0^{h,G}$  is rather significant.

We can draw a few conclusions: (a) for the given  $h$ ,  $\Psi_h(\lambda_0^h)$  is not a good approximation of  $\Psi(\lambda_0)$  (note that  $\Psi_h(\lambda_s)$  seems to be a sufficient approximation of  $\Psi(\lambda_s)$  because GR leads to a change of only 3%); (b) we can expect that  $\Psi(\lambda_0)$  is “close” to zero; (c) the values  $\Psi_h(\lambda_0^h)$ ,  $\Psi_h^G(\lambda_0^h)$ , and  $\Psi_h^G(\lambda_0^{h,G})$  are significantly different; this means that both a refinement of the mesh as well as further minimization in the neighborhood of  $\lambda_0^h$  or  $\lambda_0^{h,G}$  are necessary to gain confidence in the calculated minimum.

*Acknowledgement:* The author is indebted to Dr. Richard (Dick) Haas for many fruitful discussions.

## References

- [1] E.J. Haug, K.K. Choi, V. Komkov: *Design sensitivity analysis of structural systems*, Academic Press, Orlando, 1986.
- [2] I. Hlaváček, M. Křížek: *On a superconvergent finite element scheme for elliptic systems*, Parts I–III. *Appl. Math.* **32** (1987), 131–154, 200–213, 276–289.
- [3] I. Hlaváček, M. Křížek, V. Pištora: *How to recover the gradient of linear elements on nonuniform triangulations*, *Appl. Math.* **41** (1996), 241–267.
- [4] I. Hlaváček, M. Křížek: *Optimal error and local error estimates of a recovered gradient of linear elements on nonuniform triangulations*, *J. Comput. Math.* **14** (1996), 345–362.
- [5] M. Křížek, P. Neittaanmäki: *Superconvergence phenomenon in the finite element method arising from averaging gradients*, *Numer. Math.* **45** (1984), 105–116.
- [6] M. Křížek, P. Neittaanmäki: *On a global superconvergence of the gradient of linear triangular elements*, *J. Comput. Appl. Math.* **18** (1987), 221–233.
- [7] N. Levine: *Superconvergent recovery of the gradient from piecewise linear finite-element approximations*, *IMA J. Numer. Anal.* **5** (1985), 407–427.
- [8] O. C. Zienkiewicz, J.Z. Zhu: *A simple error estimator and adaptive procedure for practical engineering analysis*, *Int. J. Num. Meth. Eng.* **24** (1987), 337–357.