INTERNATIONAL JOURNAL OF NUMERICAL ANALYSIS AND MODELING Volume 14, Number 2, Pages 306-311

© 2017 Institute for Scientific Computing and Information

# A SIMPLE FINITE ELEMENT METHOD FOR NON-DIVERGENCE FORM ELLIPTIC EQUATIONS

LIN MU AND XIU YE

**Abstract.** We develop a simple finite element method for solving second order elliptic equations in non-divergence form by combining least squares concept with discontinuous approximations. This simple method has a symmetric and positive definite system and can be easily analyzed and implemented. Also general meshes with polytopal element and hanging node can be used in the method. We prove that our finite element solution approaches to the true solution when the mesh size approaches to zero. Numerical examples are tested that demonstrate the robustness and flexibility of the method.

Key words. Finite element methods, non-divergence form elliptic equations, polyhedral meshes.

### 1. Introduction

We consider a elliptic equations in non-divergence form

(1) 
$$A: D^2 u = f, \quad \text{in } \Omega,$$

(2) 
$$u = 0, \text{ on } \partial\Omega,$$

where  $\Omega$  is a convex polytopal domain in  $\mathbb{R}^d$  with d = 2, 3. We assume that the model problem (1)-(2) has a unique solution and the coefficient tensor A(x) is uniformly elliptic.

Non-divergence form elliptic partial differential equations have many applications in the areas such as stochastic processes and game theory [3]. In recent years, many numerical methods have been developed for second order elliptic equations in non-divergence form [1, 2, 4, 5, 6, 7] and the references therein.

The non-divergence nature of the problems makes it difficult to develop and analyze numerical algorithms for them since sophisticated Galerkin type numerical techniques cannot be applied directly. The goal of this work is to introduce a simple finite element method for non-divergence form partial differential equations which can be easily implemented and analyzed. This finite element method based on least squares methodology with discontinuous approximations has symmetric and positive definite system and is flexible to work with general meshes. We prove an optimal order error estimate for the finite element approximation in a discrete  $H^2$  norm. However, our numerical results show optimal order of convergence in a discrete  $H^1$  and  $H^2$  norm.

# 2. Finite Element Methods

Let  $\mathcal{T}_h$  be a partition of a domain  $\Omega$  consisting of polygons in two dimension or polyhedra in three dimension satisfying a set of conditions specified in [8]. Denote

Received by the editors on May 22, 2016, and accepted on October 14, 2016.

<sup>2000</sup> Mathematics Subject Classification. 35R35, 49J40, 60G40.

307

by  $\mathcal{E}_h$  the set of all edges or flat faces in  $\mathcal{T}_h$ , and let  $\mathcal{E}_h^0 = \mathcal{E}_h \setminus \partial \Omega$  be the set of all interior edges or flat faces. For every element  $T \in \mathcal{T}_h$ , we denote by  $h_T$  its diameter and mesh size  $h = \max_{T \in \mathcal{T}_h} h_T$  for  $\mathcal{T}_h$ .

We define a finite element space  $V_h$  as follows for  $k \ge 2$ ,

(3) 
$$V_h = \{ v \in L^2(\Omega) : v \in P_k(T), T \in \mathcal{T}_h \}.$$

Let elements  $T_1$  and  $T_2$  have e as a common edge/face. We define a jump of  $\phi$  on e as

$$[\phi]_e = \begin{cases} \phi|_{\partial T_1} - \phi|_{\partial T_2}, & e \in \mathcal{E}_h^0, \\ \phi, & e \in \partial \Omega. \end{cases}$$

The order of  $T_1$  and  $T_2$  is non-essential as long as the difference is taken in a consistent way.

We introduce two bilinear forms as follows

$$\begin{split} s(v,w) &= \sum_{e \in \mathcal{E}_h} \int_e h_e^s[v][w] ds + \sum_{e \in \mathcal{E}_h^0} \int_e h_e^t[\nabla v] \cdot [\nabla w] ds, \\ a(v,w) &= \sum_{T \in \mathcal{T}_h} (A:D^2v, \ A:D^2w)_T + s(v,w), \end{split}$$

where s and t are two integers such that  $s \ge -3$  and  $t \ge -1$ . For simplicity, we will let s = t = -1 in the rest of the paper.

**Algorithm 1.** A numerical approximation for (1)-(2) can be obtained by seeking  $u_h \in V_h$  satisfying the following equation:

(4) 
$$a(u_h, v) = (f, A : D^2 v) \quad \forall v \in V_h.$$

Lemma 1. The finite element scheme (4) has a unique solution.

*Proof.* It suffices to show that the solution of (4) is trivial if f = 0. Assuming f = 0 and taking  $v = u_h$  in (4), we have

$$\sum_{T \in \mathcal{T}_h} (A : D^2 u_h, A : D^2 u_h)_T + s(u_h, u_h) = 0,$$

which implies that  $A: D^2u_h = 0$  on each element T and  $u_h \in C_0^1(\Omega)$ . Thus  $u_h$  is a solution of the problem (1)-(2) with f = 0. The uniqueness of the solution of the model problem (1)-(2) implies  $u_h = 0$ .

We define a semi-norm  $|\!|\!| \cdot |\!|\!|$  as follows,

$$|||v|||^2 = a(v,v).$$

Similar to the proof of Lemma 1, we can prove that  $\|\cdot\|$  define a norm in  $V_h$ .

#### 3. Error Estimate

In this section, we will estimate the difference between the true solution u and its finite element approximation  $u_h$  from (4).

For any function  $\varphi \in H^1(T)$ , the following trace inequality holds true (see [8] for details):

(5) 
$$\|\varphi\|_e^2 \le C\left(h_T^{-1}\|\varphi\|_T^2 + h_T\|\nabla\varphi\|_T^2\right).$$

Lemma 2. The bilinear form  $a(\cdot, \cdot)$  satisfies the following continuity property, (6)  $a(v, w) \leq C |||v||| ||w||.$ 

Proof. It follows from the Cauchy-Schwarz inequality,

$$\begin{split} a(v,w) &= \sum_{T \in \mathcal{T}_h} (A : D^2 v, A : D^2 w)_T + \sum_{e \in \mathcal{E}_h} \int_e h_e^{-1}[v][w] ds \\ &+ \sum_{e \in \mathcal{E}_h^0} \int_e h_e^{-1} [\nabla v] \cdot [\nabla w] ds \\ &\leq (\sum_{T \in \mathcal{T}_h} \|A : D^2 v\|_T^2)^{\frac{1}{2}} (\sum_{T \in \mathcal{T}_h} \|A : D^2 w\|_T^2)^{\frac{1}{2}} \\ &+ (\sum_{e \in \mathcal{E}_h} \int_e h_e^{-1}[v]^2 ds)^{\frac{1}{2}} (\sum_{e \in \mathcal{E}_h} \int_e h_e^{-1}[w]^2 ds)^{\frac{1}{2}} \\ &+ (\sum_{e \in \mathcal{E}_h^0} \int_e h_e^{-1} [\nabla v]^2 ds)^{\frac{1}{2}} (\sum_{e \in \mathcal{E}_h^0} \int_e h_e^{-1} [\nabla w] 2)^{\frac{1}{2}} \\ &\leq C \| v \| \| \| w \| , \end{split}$$

which finishes the proof.

**Lemma 3.** Let u be the solution of (1)-(2) and  $Q_h u \in V_h$  be the  $L^2$  projection of u defined element by element. Then we have

(7) 
$$|||u - Q_h u||| \le Ch^{k-1} ||u||_{k+1}.$$

*Proof.* Using the definition of  $Q_h$  and (5), we have

$$\begin{aligned} \|u - Q_h u\|^2 &= \sum_{T \in \mathcal{T}_h} \|A : D^2 (u - Q_h u)\|_T^2 + \sum_{e \in \mathcal{E}_h} h_e^{-1} \|[u - Q_h u]\|_e^2 \\ &+ \sum_{e \in \mathcal{E}_h^0} h_e^{-1} \|[\nabla (u - Q_h u)]\|_e^2 \\ &\leq C \sum_{T \in \mathcal{T}_h} (|u - Q_h u|_{H^2(T)}^2 + h^{-2} \|u - Q_h u\|_T^2 + h^{-2} \|\nabla (u - Q_h u)\|_T^2) \\ &\leq C h^{2k-2} \|u\|_{k+1}^2. \end{aligned}$$

We have proved the lemma.

**Theorem 1.** Let  $u_h \in V_h$  be the finite element solution of the problem (1)-(2) arising from (4). Then there exists a constant C such that

(8) 
$$|||u - u_h||| \le Ch^{k-1} ||u||_{k+1}.$$

*Proof.* Obviously, the true solution u of (1)-(2) satisfies,

$$a(u,v) = (f, A: D^2v), \quad \forall v \in V_h.$$

Subtracting (4) from the above equation implies

$$a(u-u_h,v)=0, \quad \forall v \in V_h.$$

By adding and subtracting  $Q_h u$  in the above equation, we have

(9) 
$$a(Q_hu - u_h, v) = -a(u - Q_hu, v), \quad \forall v \in V_h.$$

308

Using (9), (6) and (7), we arrive

$$\begin{split} \|\|Q_{h}u - u_{h}\|\|^{2} &= a(Q_{h}u - u_{h}, Q_{h}u - u_{h}) \\ &= |a(u - Q_{h}u, Q_{h}u - u_{h})| \\ &\leq \|\|u - Q_{h}u\|\|\|Q_{h}u - u_{h}\|\| \\ &\leq Ch^{k-1}\|u\|_{k+1}\|Q_{h}u - u_{h}\|, \end{split}$$

which implies

(10)  $|||Q_h u - u_h||| \le Ch^{k-1} ||u||_{k+1}.$ 

Using the triangle inequality, (7) and (10), we have

$$|||u - u_h||| \le ||u - Q_h u||| + ||Q_h u - u_h||| \le Ch^{k-1} ||u||_{k+1}.$$

The proof of the theorem is completed.

### 4. Numerical Experiments

In this section, two numerical examples are tested to demonstrate the robustness of the method and to validate the theoretical conclusions in the previous section. In the following experiments, rectangular mesh and quadratic polynomial (k = 2) are used. The size of each element is denoted by h.

**4.1. Test 1.** In this numerical example, a non-divergence form elliptic equations (1)-(2) is considered with  $\Omega = (-1, 1)^2$ . Four different coefficient matrices  $A = A_i$  (i = 1, ..., 4) are investigated:

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 1 & 100 \\ 100 & 1 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 1 & 100 \\ 50 & 10 \end{pmatrix}.$$

and the exact solution is chosen as

$$u = \sin(\pi x)\sin(\pi y).$$

The error measured in  $L^2$ -norm,  $H^1$ -norm, and  $H^2$ -norm for the above four different coefficient matrices are reported in Table 1 respectively. It shows that the rate of convergence in  $H^2$  norm is optimal order  $\mathcal{O}(h)$ , which confirms our theory. Moreover, the error measured in  $H^1$ -norm converges with an optimal convergence rate of  $\mathcal{O}(h^2)$ . The rate of convergence of the  $L^2$ -error has the order  $\mathcal{O}(h^2)$ .

**4.2.** Test **2**. For this test we take A as the following continuous matrix-valued function:

$$A(x) = \begin{pmatrix} |x|^{1/2} + 1 & -|x|^{1/2} \\ -|x|^{1/2} & 5|x|^{1/2} + 1 \end{pmatrix}.$$

Let  $\Omega = (-1/2, 1/2)^2$  and choose f such that the exact solution is given by

 $u(x, y) = \sin(2\pi x)\sin(2\pi y)\exp(x\cos(y)).$ 

It is easy to see that the problem is coupled with homogeneous boundary condition.

The error measured in  $L^2$ -norm,  $H^1$ -norm, and  $H^2$ -norm are reported in Table 2. Again the error measured in the  $H^2$ -norm is of order  $\mathcal{O}(h)$  as expected. The errors measured in  $H^1$ -norm and  $L^2$ -norm converge at the order  $\mathcal{O}(h^2)$ .

## L. MU AND X. YE

TABLE 1. Example 1. Convergence test on rectangular mesh for k = 2.

1/h	$\ u-u_h\ $	Rate	$\ \nabla u - \nabla u_h\ $	Rate	$\ D^2u - D^2u_h\ $	Rate			
	$A = A_1$								
2	1.4952E-01		9.4428E-01		$5.7801E{+}00$				
4	4.6597 E-02	1.68	2.6559E-01	1.83	$3.0925E{+}00$	0.90			
8	1.2618E-02	1.88	6.6617 E-02	2.00	$1.5732E{+}00$	0.98			
16	3.2025E-03	1.98	1.6675 E-02	2.00	7.9000E-01	0.99			
32	8.0271E-04	2.00	4.1624 E-03	2.00	3.9543E-01	1.00			
64	2.0077 E-04	2.00	1.0389E-03	2.00	1.9777E-01	1.00			
128	5.0198 E-05	2.00	2.5946E-04	2.00	9.8890 E-02	1.00			
	$A = A_2$								
2	1.5229E-01		8.8746E-01		8.1743E + 00				
4	7.0955E-02	1.10	3.7333E-01	1.25	$4.3735E{+}00$	0.90			
8	2.3711E-02	1.58	1.1844E-01	1.66	2.2248E + 00	0.98			
16	7.2493E-03	1.71	3.6778 E-02	1.69	1.1172E + 00	0.99			
32	2.1085E-03	1.78	1.1180E-02	1.72	5.5922E-01	1.00			
64	5.8423E-04	1.85	3.2542E-03	1.78	2.7969E-01	1.00			
128	1.5567E-04	1.91	9.0403 E-04	1.85	1.3985 E-01	1.00			
	$A = A_3$								
2	1.2647E-01		6.8922 E-01		5.7804E + 02				
4	4.0909E-02	1.63	2.2186E-01	1.64	3.0927E + 02	0.90			
8	1.1112E-02	1.88	5.9639E-02	1.90	1.5732E + 02	0.98			
16	2.8509E-03	1.96	1.5228E-02	1.97	7.9004E + 01	0.99			
32	7.1948E-04	1.99	3.8336E-03	1.99	$3.9545E{+}01$	1.00			
64	1.8066E-04	1.99	9.6125 E-04	2.00	1.9778E + 01	1.00			
128	4.5301E-05	2.00	2.4080E-04	2.00	9.8895E + 00	1.00			
	$A = A_4$								
2	1.2839E-01		6.9736E-01		4.3467E + 02				
4	4.2163E-02	1.61	2.2787 E-01	1.61	2.3256E + 02	0.90			
8	1.1654 E-02	1.86	6.2194 E-02	1.87	1.1830E + 02	0.98			
16	3.0432E-03	1.94	1.6091 E-02	1.95	$5.9409E{+}01$	0.99			
32	7.7716E-04	1.97	4.0870E-03	1.98	$2.9737E{+}01$	1.00			
64	1.9680E-04	1.98	1.0331E-03	1.98	$1.4872E{+}01$	1.00			
128	4.9736E-05	1.98	2.6138E-04	1.98	7.4367E + 00	1.00			

TABLE 2. Example 2. Convergence test on rectangular mesh for k = 2.

1/h	$\ u-u_h\ $	Rate	$\ \nabla u - \nabla u_h\ $	Rate	$\ D^2u - D^2u_h\ $	Rate
2	2.0604E-01		2.7963E+00		6.8799E + 01	
4	1.8951E-01	0.12	2.1934E + 00	0.35	$6.3871E{+}01$	0.11
8	6.5922E-02	1.52	6.8159E-01	1.69	3.4699E + 01	0.88
16	1.8674 E-02	1.82	1.8351E-01	1.89	$1.7719E{+}01$	0.97
32	4.8056 E-03	1.96	4.6848E-02	1.97	$8.9097E{+}00$	0.99
64	1.1913E-03	2.01	1.1738E-02	2.00	4.4630E + 00	1.00
128	2.9070E-04	2.03	2.9413 E-03	2.00	2.2334E + 00	1.00

#### Acknowledgments

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research. This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan). The work of X. Ye is supported in part by the National Science Foundation under contract DMS-1620016.

### References

- A. Dedner and T. Pryer. Discontinuous Galerkin methods for non variational problems, arXiv:1304.2265v1.
- [2] X. Feng, L. Hennings, and M. Neilan. Finite element methods for second order linear elliptic partial differential equations in non divergence form, Math. Comp., to appear.
- [3] W. Fleming and H. Soner, Controlled Markov Processes and Viscosity Solutions, 2nd ed., Stoch. Model. Appl. Probab. 25, Springer, New York, 2006.
- [4] O. Lakkis, and T. Pryer. A finite element method for second order nonvariational elliptic problems, SIAM J. Sci. Comput., 33 (2011), 786-801.
- [5] R.H. Nochetto and W. Zhang. Discrete ABP estimate and convergence rates for linear elliptic equations in nondivergence form, arXiv:1411.6036.
- [6] I. Smears and E. Suli. Discontinuous Galerkin finite element approximation of non-divergence form elliptic equations with Cord'es coefficients, SIAM J. Numer. Anal., 51 (2013), 2088-2106.
- [7] C. Wang and J. Wang. A primal-dual weak Galerkin finite element method for second order elliptic equations in non-divergence form. arXiv:1510.03499.
- [8] J. Wang and X. Ye, A Weak Galerkin mixed finite element method for second-order elliptic problems, Mathematics of Computation, 83 (2014), 2101-2126. arXiv:1202.3655v1.

Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, United States

*E-mail*: mul1@ornl.gov

Department of Mathematics, University of Arkansas at Little Rock, Little Rock, AR 72204, United States

*E-mail*: xxye@ualr.edu