ON THE SOLUTION OF A CLASS OF ELLIPTIC PROBLEMS

Davod Khojasteh Salkuyeh

Department of Mathematics, Mohaghegh Ardabili University,
P. O. Box. 56199-11367, Ardabil, Iran
E-mail: khojaste@uma.ac.ir

Abstract

In this paper, a class of elliptic partial differential equations (PDEs) is considered. A sufficient condition for the existence and uniqueness of a solution to the problem is obtained. Discretizing the equation by the finite difference method yields a linear system whose matrix is nonsymmetric and banded. We show that under a certain condition the discretization parameter can be chosen small enough so that the linear system has a unique solution. An algorithm is presented to build the matrix and the vector on the right-hand side of the linear system. The numerical tests performed confirm the efficiency of the algorithm.

Mathematics Subject Classification: 65F.

Keywords: elliptic problem, finite difference, Lax-Milgram theorem.

1 Introduction

Consider the following elliptic PDE

\[-\frac{\partial}{\partial x}(p \frac{\partial u}{\partial x}) - \frac{\partial}{\partial y}(q \frac{\partial u}{\partial y}) + \text{div}(ru) + gu = f, \quad x \in \Omega = (a, b) \times (c, d), \quad (1a)\]

\[u = 0 \quad \text{on} \quad \Gamma = \partial \Omega, \quad (1b)\]

where

\[x = (x, y), \quad r = (d, e), \quad d, e, p, q \in C^1(\Omega), \quad g \in C(\Omega), \quad f \in L^2(\Omega),\]

and for some \(\rho > 0\)

\[p(x, y), q(x, y) \geq \rho, \quad \forall (x, y) \in \Omega.\]
In the case that \( p = q = 1 \), the problem has been studied in [3]. Let \( v \) be in \( H^2_0(\Omega) \), then by multiplying both sides of (1a) by \( v \) and integrating over \( \Omega \) and using Green’s theorem, the following weak formulation of (1a) and (1b) is obtained:

Find \( u \in H^1_0(\Omega) \) such that

\[
B(u, v) = F(v), \quad \forall v \in H^1_0(\Omega),
\]

where the bilinear form \( B \) and the functional \( F \) are defined by

\[
B(u, v) = \int_{\Omega} \left( p \frac{\partial u}{\partial x} \frac{\partial v}{\partial x} + q \frac{\partial u}{\partial y} \frac{\partial v}{\partial y} - ru. \nabla v + guv \right) dx, \quad F(v) = \int_{\Omega} f v dx.
\]

The solution \( u \) of (2) is said to be the weak solution of (1a) and (1b). If \( u \) is sufficiently smooth, then the weak solution \( u \) is also the classical solution of (1a) and (1b) [3].

**Theorem 1.1** If for some \( \alpha > 0 \), we have

\[
\frac{1}{2} \text{div } r + g \geq \alpha, \quad \forall x \in \Omega,
\]

then, the problem (2) has a unique solution.

**Proof.** If we can show that \( F \) is continuous and that \( B \) is both continuous and coercive, then by the Lax-Milgram theorem [1, 5], we are guaranteed the existence of a unique solution to (2). First, \( F \) is continuous since

\[
|F(v)| = |\int_{\Omega} f v dx| \leq |(f, v)|_{L^2(\Omega)} \leq \|f\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \leq k \|v\|_{H^1(\Omega)},
\]

where \( k = \|f\|_{H^2(\Omega)} \).

Next, \( B \) is continuous since

\[
|B(u, v)| \leq \gamma \int_{\Omega} |\nabla u| \nabla v|dx + \int_{\Omega} |ru. \nabla v| dx + \int_{\Omega} |guv| dx
\]

\[
(\gamma = \max\{\sup_{x \in \Omega} p(x), \sup_{x \in \Omega} q(x)\})
\]

\[
\leq \gamma \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)} + \sup_{x \in \Omega} |r| \|u\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} + \sup_{x \in \Omega} |g(x)| \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)}
\]

\[
(\text{note that } \|w_1, w_2\|_{L^2(\Omega)}^2 = \int_{\Omega} (w_1^2 + w_2^2) dx)
\]

\[
\leq \gamma \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)} + \sup_{x \in \Omega} |r| \|u\|_{L^2(\Omega)} \|v\|_{H^1(\Omega)} + \sup_{x \in \Omega} |g(x)| \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)}
\]

\[
\leq (\gamma + \sup_{x \in \Omega} |r| + \sup_{x \in \Omega} |g(x)|) \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)}.
\]
Hence 
\[ |B(u, v)| \leq C \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)}, \]
where 
\[ C = (\gamma + \sup_{x \in \Omega} |r| + \sup_{x \in \Omega} |g(x)|). \]

Thus \( B \) is continuous. Now we show that \( B \) is coercive. Using integration by parts, we have
\[
\int_{\Omega} -rv \cdot \nabla v \, dx = \int_{\Omega} (-dv \frac{\partial v}{\partial x} - ev \frac{\partial v}{\partial y}) \, dx = \int_{\Omega} [v \frac{\partial}{\partial x}(dv) + v \frac{\partial}{\partial y}(ev)] \, dx
\]
\[
= \int_{\Omega} \left( \frac{\partial d}{\partial x} + \frac{\partial e}{\partial y} \right) v^2 \, dx + \int_{\Omega} rv \cdot \nabla v \, dx.
\]
Therefore
\[
\int_{\Omega} -rv \cdot \nabla v \, dx = \frac{1}{2} \int_{\Omega} \left( \frac{\partial d}{\partial x} + \frac{\partial e}{\partial y} \right) v^2 \, dx + \frac{1}{2} \int_{\Omega} (\text{div } r) v^2 \, dx.
\]

Using (4) it follows that
\[
B(v, v) = \int_{\Omega} \left[ p(x)\left( \frac{\partial u}{\partial x} \right)^2 + q(x)\left( \frac{\partial u}{\partial y} \right)^2 + \left( \frac{1}{2} \text{div } r + g \right) v^2 \right] \, dx
\]
\[
\geq \int_{\Omega} \left[ \rho \left| \nabla v \right| ^2 + \alpha v^2 \right] \, dx \geq \min\{\rho, \alpha\} \int_{\Omega} \left| \nabla v \right| ^2 + v^2 \, dx
\]
\[
= C \parallel v \parallel^2_{H^1(\Omega)},
\]
where \( C = \min\{\rho, \alpha\} \). This shows that \( B \) is coercive. All requirements are met and so a unique solution to (2) exists and the proof is completed.

Remark 1. If in the problem (1a) and (1b) the boundary condition is nonhomogeneous, it can be converted to an equivalent problem with homogeneous boundary condition, and then Theorem 1 can be applied [3, 4].

2 Approximate Problem

Let \( n_x \) and \( n_y \) be natural numbers and \( h = \frac{b-a}{n_x}, k = \frac{d-c}{n_y} \), and for simplicity we let \( h = k \). Let \( u_{i,j} \) be the approximate value of \( u \) at the grid point \((x_j, y_i)\) where \( x_j = a + jh, y_i = c + ih \). Consider the following centered difference formula ([8])
\[
-\frac{\partial}{\partial x}(p \frac{\partial u}{\partial x}) - \frac{\partial}{\partial y}(q \frac{\partial u}{\partial y}) = \frac{p_{i,j+\frac{1}{2}}(u_{i,j+1} - u_{i,j}) - p_{i,j-\frac{1}{2}}(u_{i,j} - u_{i,j-1})}{h^2}
\]
\[
+ \frac{q_{i+\frac{1}{2},j}(u_{i+1,j} - u_{i,j}) - q_{i-\frac{1}{2},j}(u_{i,j} - u_{i-1,j})}{h^2} + O(h^2).
\]
Discretizing the equation (1a) by using the above centered difference formula yields
\[
\begin{align*}
\frac{-1}{h^2} \left[q_{i+\frac{1}{2},j} u_{i+1,j} + q_{i-\frac{1}{2},j} u_{i-1,j} + p_{i,j+\frac{1}{2}} u_{i,j+1} + p_{i,j-\frac{1}{2}} u_{i,j-1}
\right. \\
\left. - \left(q_{i+\frac{3}{2},j} + q_{i-\frac{3}{2},j} + p_{i,j+\frac{3}{2}} + p_{i,j-\frac{3}{2}}\right) u_{i,j}\right]
\end{align*}
\]
\[
\frac{1}{2h} \left[d_{i,j+1} u_{i,j+1} - d_{i,j-1} u_{i,j-1}\right] + \frac{1}{2h} \left[e_{i+1,j} u_{i+1,j} - e_{i-1,j} u_{i-1,j}\right] + g_{i,j} u_{i,j} = f_{i,j}
\]
for \(i = 1, 2, \ldots, n_y - 1\), \(j = 1, 2, \ldots, n_x - 1\), or
\[
\begin{align*}
-q_{i+\frac{1}{2},j} - \frac{h}{2} e_{i+1,j} u_{i+1,j} - q_{i-\frac{1}{2},j} - \frac{h}{2} e_{i-1,j} u_{i-1,j} - p_{i,j+\frac{1}{2}} - \frac{h}{2} d_{i,j+1} u_{i,j+1}

-p_{i,j-\frac{1}{2}} + \frac{h}{2} d_{i,j-1} u_{i,j-1} + q_{i+\frac{1}{2},j} + q_{i-\frac{1}{2},j} + p_{i,j+\frac{1}{2}} + p_{i,j-\frac{1}{2}} + h^2 g_{i,j} u_{i,j} = h^2 f_{i,j}. \quad (5)
\end{align*}
\]
Note that the truncation error for (5) is \(O(h^2)\).

The equations in (5) may be written in matrix form as
\[
A \mathbf{x} = \mathbf{b}, \quad (6)
\]
where the vectors \(\mathbf{b}\) and \(\mathbf{x}\) are in the following forms
\[
\mathbf{b} = h^2 \left[f_{1,1}, \ldots, f_{1,n_x-1}; f_{2,1}, \ldots, f_{2,n_x-1}; \ldots; f_{n_y-1,1}, \ldots, f_{n_y-1,n_x-1}\right]^T,
\]
\[
\mathbf{x} = [u_{1,1}, \ldots, u_{1,n_x-1}; u_{2,1}, \ldots, u_{2,n_x-1}; \ldots u_{n_y-1,1}, \ldots, u_{n_y-1,n_x-1}]^T.
\]
It can be easily shown that \(A\) is a band matrix with half bandwidth \(n_x - 1\) in the following form
\[
A = \\
\begin{bmatrix}
E_1 & F_1 & H_1 \\
D_2 & E_2 & F_2 & H_2 \\
& D_3 & E_2 & F_2 & H_3 \\
& & \ddots & \ddots & \ddots \\
B_{n_x} & & & \ddots & \ddots & \ddots \\
B_{n_x+1} & \ddots & \ddots & \ddots & \ddots & \ddots \\
& \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
& & \ddots & \ddots & \ddots & \ddots & \ddots \\
& & & \ddots & \ddots & \ddots & \ddots & \ddots \\
& & & & \ddots & \ddots & \ddots & \ddots \\
B_n & & & & & \ddots & \ddots & \ddots \\
D_n & & & & & & F_{n-1} & E_n
\end{bmatrix}
\]

**Theorem 2.1** If
\[
\text{div} \mathbf{r} + g > 0, \quad \forall \mathbf{x} \in \Omega, \quad (7)
\]
then the system (6) has a unique solution provided \(h\) is chosen sufficiently small.
Proof. The entries in each row of $A$ are in the following forms

$$
\beta_1 = -\left(q_{i+\frac{1}{2},j} - \frac{h}{2}e_{i+1,j}\right), \quad \beta_2 = -\left(q_{i-\frac{1}{2},j} + \frac{h}{2}e_{i-1,j}\right), \quad \beta_3 = -\left(p_{i,j+\frac{1}{2}} + \frac{h}{2}d_{i,j+1}\right)
$$

$$
\beta_4 = -\left(p_{i,j-\frac{1}{2}} + \frac{h}{2}d_{i,j-1}\right), \quad \beta_0 = q_{i+\frac{1}{2},j} + q_{i-\frac{1}{2},j} + p_{i,j+\frac{1}{2}} + p_{i,j-\frac{1}{2}} + h^2g_{i,j},
$$

where $\beta_0$ is a diagonal entry. Now for $h$ sufficiently small we have

$$
\beta_0 > 0, \quad \beta_i < 0, \quad i = 1, 2, 3, 4. \quad (8)
$$

We show that $h$ can be chosen so that $A$ is strictly diagonally dominant. The condition (7) can be written as

$$
\frac{\partial d}{\partial x} + \frac{\partial e}{\partial y} + g > 0, \quad (x, y) \in \Omega.
$$

Thus, for $h$ sufficiently small we have

$$
\frac{d_{i,j+1} - d_{i,j-1}}{2h} + \frac{e_{i+1,j} - e_{i-1,j}}{2h} + g_{ij} > 0, \quad (x, y) \in \Omega, \quad (9)
$$

or

$$
q_{i+\frac{1}{2},j} + q_{i-\frac{1}{2},j} + p_{i,j+\frac{1}{2}} + p_{i,j-\frac{1}{2}} + h^2g_{ij} > \left(q_{i+\frac{1}{2},j} - \frac{h}{2}e_{i+1,j}\right)
$$

$$
+ \left(q_{i-\frac{1}{2},j} + \frac{h}{2}e_{i-1,j}\right) + \left(p_{i,j+\frac{1}{2}} - \frac{h}{2}d_{i,j+1}\right) + \left(p_{i,j-\frac{1}{2}} + \frac{h}{2}d_{i,j-1}\right).
$$

Hence

$$
\left| \beta_0 \right| > \sum_{i=1}^{4} \left| \beta_i \right|.
$$

This shows that $A$ is strictly diagonally dominant, and so is invertible. Therefore, if $h$ is chosen so that the relations (8) and (9) hold simultaneously, $A$ is invertible.

Remark 2. If the condition (7) is replaced by

$$
\frac{d_{i,j+1} - d_{i,j-1}}{2h} + \frac{e_{i+1,j} - e_{i-1,j}}{2h} + g_{ij} \geq 0, \quad \forall \ i, j
$$

for $h$ sufficiently small, then

$$
\left| \beta_0 \right| \geq \sum_{i=1}^{4} \left| \beta_i \right| \quad (10)
$$

and in this case it can be easily shown that for the first and the last rows of $A$, the inequality in (10) is strict. Since $\beta_i \neq 0$, for $i=0,1,2,3,4$, $A$ is irreducible. Therefore $A$ is irreducibly diagonally dominant and hence is invertible [6].
3 Algorithm

Using the following algorithm, we may form the matrix $A$ and the vector $b$ in (6). Since this matrix has five nonzero diagonals, it suffices to store these diagonals in five vectors as follows:

$$B = [0, \ldots, 0, B_{n_x}, \ldots, B_n]^T,$$
$$D = [0, D_2, \ldots, D_n]^T,$$
$$E = [E_1, \ldots, E_n]^T,$$
$$F = [F_1, \ldots, F_{n-1}, 0]^T,$$
$$H = [H_1, \ldots, H_{n-n_x+1}, 0, \ldots, 0]^T.$$

In the algorithm, $u_d, u_r, u_u,$ and $u_l$ denote the boundary conditions on the lines $y = c, x = b, y = d,$ and $x = a,$ respectively.

**Algorithm**

\[
\begin{align*}
n_x &:= n_x - 1 \\
n_y &:= n_y - 1 \\
n &:= n_x \times n_y \\
\text{For } m = 1, \ldots, n \text{ Do} \\
\quad j_{\text{temp}} &:= \left\lfloor \frac{m}{n_x} \right\rfloor \\
\quad i_{\text{temp}} &:= m - n_x \times j_{\text{temp}} \\
\quad &\text{If } i_{\text{temp}} = 0 \text{ then} \\
\quad &\quad j := n_x \\
\quad &\quad i := j_{\text{temp}} \\
\quad &\text{Else} \\
\quad &\quad j := i_{\text{temp}} \\
\quad &\quad j := j_{\text{temp}} + 1 \\
\quad \text{EndIf} \\
\quad p_1 &:= p_{i,j} + \frac{1}{2} \\
\quad p_2 &:= p_{i,j} - \frac{1}{2} \\
\quad q_1 &:= q_{i+rac{1}{2},j} \\
\quad q_2 &:= q_{i-rac{1}{2},j} \\
\quad T &:= p_1 + p_2 + q_1 + q_2 \\
\quad E(m) &:= E(m) + \frac{1}{h^2} T + g_{ij} \\
\quad &\text{If } (m + n_x \leq n) \text{ then} \\
\quad &\quad H(m) := H(m) - \frac{1}{h^2} q_1 + \frac{1}{2h} e_{i+1,j} \\
\quad &\text{EndIf} \\
\quad &\text{If } (m - n_x \geq 1) \text{ then} \\
\quad &\quad B(m) := B(m) - \frac{1}{h^2} q_2 - \frac{1}{2h} e_{i-1,j} \\
\quad &\text{EndIf} \\
\quad &\text{If } (j + 1 \leq n_x) \text{ then}
\end{align*}
\]
On The Solution of Class of Elliptic Problems

4 Numerical examples

Example 1. Consider the equation

\[ -\frac{\partial}{\partial x} (x^2y \frac{\partial u}{\partial x}) - \frac{\partial}{\partial y} (xy^2 \frac{\partial u}{\partial y}) + \frac{\partial}{\partial x} ((x - y)u) + \frac{\partial}{\partial y} ((x + y)u) \]

\[ + u = 5(x^2 + y^2) - 6xy(x + y), \quad (x, y) \in \Omega = (1, 2) \times (1, 3), \]

with the following boundary conditions

\[ u(x, y) = \begin{cases} 
  x^2 + 1, & 1 \leq x \leq 2, \quad y = 1, \\
  y^2 + 4, & 1 \leq y \leq 3, \quad x = 2, \\
  x^2 + 9, & 1 \leq x \leq 2, \quad y = 3, \\
  y^2 + 1, & 1 \leq y \leq 3, \quad x = 1.
\]

In this example we have

\[ p(x, y) = x^2y \geq 1 = \rho \quad \text{and} \quad q(x, y) = xy^2 \geq 1 = \rho, \]
and

\[ d(x, y) = x - y, \quad e(x, y) = x + y, \quad g(x, y) = 1, \quad f(x, y) = 5(x^2 + y^2) - 6xy(x + y). \]

Now, we see that

\[ \text{div } \mathbf{r} + g = \frac{\partial d}{\partial x} + \frac{\partial e}{\partial y} + g = 3 > 0, \quad \frac{1}{2} \text{div } \mathbf{r} + g = 2 > 0. \]

So, by Theorem 1.1 the existence of a unique solution to the above problem is guaranteed. Also by Theorem 2.1, the approximate solution at each grid point exists and it is unique. The exact solution for this problem is \( u(x, y) = x^2 + y^2 \).

In Figure 1, the graph of the absolute value of error has been shown at the grid points. The approximate solution has been obtained using the algorithm presented in Section 3. In this example, we take \( h = 0.04 \), so that \( n = 1176 \). The GMRES(m) method [6, 7] with \( m = 15 \) has been used for solving the linear system. The stopping criterion used is that the 2-norm of the residual is at most \( 10^{-6} \). The results has been obtained after 45 iterations.

**Example 2.** Consider the equation

\[ -\frac{\partial}{\partial x}(e^x \frac{\partial u}{\partial x}) - \frac{\partial}{\partial y}(e^y \frac{\partial u}{\partial y}) + \frac{\partial}{\partial x}(e^x u) - xu = xe^x + e^{-x} - x(x + ye^{-(x+y)}), \]

\((x, y) \in \Omega = (0, 1) \times (0, 1)\),
with the following boundary conditions

\[ u(x, y) = \begin{cases} 
  x, & 0 \leq x \leq 1, \ y = 0, \\
  1 + ye^{-(1+y)}, & 0 \leq y \leq 1, \ x = 1, \\
  x + e^{-(x+1)}, & 0 \leq x \leq 1, \ y = 1, \\
  ye^{-y}, & 0 \leq y \leq 1, \ x = 0,
\end{cases} \]

Here we have

\[ p(x, y) = e^x \geq 1 = \rho \quad \text{and} \quad q(x, y) = e^y \geq 1 = \rho, \]

and

\[ d(x, y) = e^x, \ e(x, y) = 0, \ g(x, y) = -x, \ f(x, y) = xe^x + e^{-x} - x(x + ye^{-(x+y)}). \]

Now we see that

\[ \text{div} \ r + g = e^x - x \geq 1 > 0, \quad \frac{1}{2} \text{div} \ r + g = \frac{1}{2}e^x - x \geq 1 - \ln 2 > 0. \]

So, by theorems 1.1 and 2.1 the existence of a unique solution to the above problem and its approximation at each grid point are guaranteed. In this example the exact solution is \( u(x, y) = x + ye^{-(x+y)} \). Here \( h = 0.025, n = 1521, m = 15 \), and the stopping criterion is the same as in example 1. In this example, the results has been obtained after 27 iterations. The absolute error has been shown in Figure 2.
References


Received: August 31, 2005