ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ
ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ
ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ
"ΘΕΩΡΗΤΙΚΗ ΠΛΗΡΟΦΟΡΙΚΗ ΚΑΙ ΘΕΩΡΙΑ ΣΥΣΤΗΜΑΤΩΝ ΕΛΕΓΧΟΥ"

ΑΡΙΘΜΗΤΙΚΕΣ ΜΕΘΟΔΟΙ ΕΠΙΛΥΣΗΣ ΥΨΗΛΗΣ ΤΑΞΗΣ PDE ΜΕ ΥΨΗΛΗΣ ΤΑΞΗΣ ΜΕΘΟΔΟΥΣ - ΕΦΑΡΜΟΓΕΣ ΣΕ MATLAB

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Καρτσιώτης Γιώργος

Επιβλέπουσα : Μαρία Γουσίδου - Κουτίτα
Αν. Καθηγήτρια Α.Π.Θ.

Θεσσαλονίκη, Μάρτιος 2013

ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ
ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ
ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ
"ΘΕΩΡΗΤΙΚΗ ΠΛΗΡΟΦΟΡΙΚΗ ΚΑΙ ΘΕΩΡΙΑ ΣΥΣΤΗΜΑΤΩΝ ΕΛΕΓΧΟΥ"

ΑΡΙΘΜΗΤΙΚΕΣ ΜΕΘΟΔΟΙ ΕΠΙΛΥΣΗΣ ΥΨΗΛΗΣ ΤΑΞΗΣ PDE ΜΕ ΥΨΗΛΗΣ ΤΑΞΗΣ ΜΕΘΟΔΟΥΣ - ΕΦΑΡΜΟΓΕΣ ΣΕ MATLAB

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Καρτσιώτης Γιώργος

Επιβλέπουσα : Μαρία Γουσίδου - Κουτίτα
Αν. Καθηγήτρια Α.Π.Θ.

Θεσσαλονίκη, Μάρτιος 2013

ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ
ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ
ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ
"ΘΕΩΡΗΤΙΚΗ ΠΛΗΡΟΦΟΡΙΚΗ ΚΑΙ ΘΕΩΡΙΑ ΣΥΣΤΗΜΑΤΩΝ ΕΛΕΓΧΟΥ"

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Καρτσιώτης Γιώργος

Επιβλέπουσα : Μαρία Γουσίδου - Κουτίτα
Αν. Καθηγήτρια Α.Π.Θ.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την Μαρτίου 2013.

....	....	....
Μ. Γουσίδου	Γ. Ραχώνης	Ν. Καραμπετάκης
Αν. Καθηγητής Α.Π.Θ.	Επ. Καθηγητής Α.Π.Θ.	Αν. Καθηγητής Α.Π.Θ.

Θεσσαλονίκη, Μάρτιος 2013

………………………………………….

Καρτσιώτης Γιώργος

Πτυχιούχος - Μαθηματικό - Α.Π.Θ

Πτυχιούχος - Εφαρμοσμένη Πληροφορική - Πανεπιστήμιο Μακεδονίας

Διπλωματούχος - Προηγμένα Συστήματα Υπολογιστών και Επικοινωνιών Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών - Α.Π.Θ

ΠΕΡΙΛΗΨΗ

Η μέθοδος των πεπερασμένων στοιχείων είναι μια καλώς θεμελιωμένη και ευρύτατα εφαρμοσμένη μεθοδολογία η οποία έχει τις ρίζες της στην μαθηματική ανάλυση και οι εφαρμογές της καλύπτουν ένα τεράστιο φάσμα από στατική, μηχανική και υδροδυναμική έως και κατασκευή τρισδιάστατων μοντέλων σε ηλεκτρονικούς υπολογιστές.

Σκοπός της εργασίας τούτης είναι η διερεύνηση του θεωρητικού υποβάθρου και η υλοποίηση της μεθοδολογίας των πεπερασμένων στοιχείων με υψηλής τάξης στοιχεία σε διαφορικές εξισώσεις με μερικές παραγώγους και η διερεύνηση της αποτελεσματικότητας τους με εφαρμογές σε διαφορετικές συναρτήσεις και πεδία ορισμού.

Η διπλωματική εργασία χωρίζεται σε έξι κεφάλαια, όπου το πρώτο κεφάλαιο είναι η εισαγωγή, στο δεύτερο κεφάλαιο παρατίθεται το θεωρητικό υπόβαθρο, όπου παρουσιάζουμε τα τμήματα της θεωρίας που χρειαζόμαστε από τον τομέα της μαθηματικής ανάλυσης, στο τρίτο κεφάλαιο αναλύουμε την εφαρμογή της μεθόδου Ritz σε δυο οικογένειες εξισώσεων, στην $\frac{\partial}{\partial x}\left(p(x,y)\frac{\partial u}{\partial x}\right)+\frac{\partial}{\partial y}\left(q(x,y)\frac{\partial u}{\partial y}\right)+r(x,y)(x,y)=f(x,y)$ και την $-\frac{\partial}{\partial x}\left(p(x)\cdot\frac{\partial y}{\partial x}\right)+q(x)\cdot y(x)=f(x,y)$, και στο τέταρτο κεφάλαιο παραθέτουμε την εφαρμογή της μεθόδου Galerkin αρχικά στην γενική περίπτωση και μετέπειτα στη μια και στις δυο διαστάσεις για τις εξισώσεις Poisson και Laplace.

Στο πέμπτο κεφάλαιο συγκεντρώσαμε όλη τη μεθοδολογία και τα αποτελέσματα των πειραμάτων που αναπτύχθηκαν στα προηγούμενα κεφάλαια και αναλύουμε την αποτελεσματικότητα τους και τέλος στο έκτο κεφάλαιο παραθέτουμε όλους τους κώδικες που γραφτήκαν και χρησιμοποιήθηκαν στην Matlab με πλήρη σχολιασμό σε κάθε τμήμα τους και στο έβδομο κεφάλαιο βρίσκονται όλες οι διαδικτυακές και βιβλιογραφικές αναφορές.

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ

Πεπερασμένα στοιχεία, μερικές διαφορικές εξισώσεις, μέθοδος Ritz, μέθοδος Galerkin, χώροι Sobolev

ABSTRACT

The method of abstract finite elements is a well posed and widely applied method which has it's roots in the field of mathematical analysis and it's applications encompass a wide field from static analysis, hydrodynamics up to three dimension computer model creation on modern personal computers.

This thesis aims at developing and gathering all the mathematical concepts used by the field of mathematical analysis and how they are applied in the frame of finite element analysis on equations with partial differential equations of higher degree and is divided to six chapters; the first one is the introduction; on the second we develop the concepts we used from the mathematical analysis field;

On the third we applied at a theoretical and at a practical level the Ritz method on the equations $\frac{\partial}{\partial x}\left(p(x,y)\frac{\partial u}{\partial x}\right)+\frac{\partial}{\partial y}\left(q(x,y)\frac{\partial u}{\partial y}\right)+r(x,y)(x,y)=f(x,y)$ and $-\frac{\partial}{\partial x}\left(p(x)\cdot\frac{\partial y}{\partial x}\right)+q(x)\cdot y(x)=f(x,y)$; on the fourth chapter we applied the Galerkin method on various problems in one and two dimensions with emphasis on the Poisson and Laplace equations.

On the fifth chapter we present the methodology deployed at various points on and analyze their performance and fine points and on the sixth chapter we present the m files used with various remarks explaining their utility at all points; At the end we gathered the bibliography and web references used.

KEY WORDS

Finite element methods, partial differential equations, Ritz method, Galerkin method, Sobolev spaces.

Αφιερώνεται στην Μεταξία

Πίνακας Περιεχομένων

1. Εισαγωγή

Σκοπός της εργασίας τούτης είναι η διερεύνηση του θεωρητικού υποβάθρου και η υλοποίηση της μεθοδολογίας των πεπερασμένων στοιχείων με υψηλής τάξης στοχεία σε διαφορικές εξισώσεις με μερικές παραγώγους και η διερεύνηση της αποτελεσματικότητας τους με εφαρμογές σε διαφορετικές συναρτήσεις και πεδία ορισμού. Στο δεύτερο κεφάλαιο παρατίθεται το θεωρητικό υπόβαθρο, όπου παρουσιάζουμε τα τμήματα της θεωρίας που χρειαζόμαστε από τον τομέα της μαθηματικής ανάλυσης, αρχικά με τους χώρους $C^m(\Omega)$ και $L^m(\Omega)$ και μετέπειτα με τους χώρους Sobolev, το θεώρημα ενσωμάτωσης και την έννοια της ασθενής παραγώγου. Βάσει των παραπάνω εννοιών ορίζουμε το συμμετρικό και το ασύμμετρο πρόβλημα μεταβλητότητας τα οποία μας οδηγούν στις μεθόδους Ritz και Galerkin αντίστοιχα και μετέπειτα παραθέτουμε τους τύπους των πεπερασμένων στοιχείων που χρησιμοποιούνται στις μεθόδους αλλά και την διασύνδεση όλων των παραπάνω εννοιών με τις PDE.

Στο τρίτο κεφάλαιο αναλύουμε την εφαρμογή της μεθόδου Ritz σε δυο οικογένειες εξισώσεων, όπου για την $\frac{\partial}{\partial x}\left(p(x,y)\frac{\partial u}{\partial x}\right)+\frac{\partial}{\partial y}\left(q(x,y)\frac{\partial u}{\partial y}\right)+r(x,y)(x,y)=f(x,y)$ προχωρούμε σε πλήρη θεωρητική ανάλυση και για την $-\frac{\partial}{\partial x}\left(p(x)\cdot\frac{\partial y}{\partial x}\right)+q(x)\cdot y(x)=f(x,y)$ πέραν της ανάλυσης λύνουμε πλήρως το σύστημα με την ανάλογη υλοποίηση στη Matlab. Στο τέλος του κεφαλαίου συγκεντρώσαμε τα πειραματικά αποτελέσματα της εφαρμογής του αναδρομικού αλγορίθμου που εφαρμόστηκε στην δεύτερη εξίσωση, ο οποίος επανεκτελούταν έως ότου επιτευχθεί η επιθυμητή ακρίβεια και παρουσιάζουμε τα σχετικά γραφήματα.

Στο τέταρτο κεφάλαιο παραθέτουμε την εφαρμογή της μεθόδου Galerkin αρχικά στην γενική περίπτωση και μετέπειτα στη μια και στις δυο διαστάσεις για τις εξισώσεις Poisson και Laplace, αναλύοντας πλήρως την εφαρμογή της μεθόδου σε θεωρητικό και υπολογιστικό επίπεδο, όπου δόθηκε ιδιαίτερη μνεία στην εφαρμογή σε πολλάπλες τιμές βημάτων και στην παραγώγη σε κάθε περίπτωση του σχετικού χρωματικού κώδικα. Στο πέμπτο κεφάλαιο συγκεντρώσαμε όλη τη μεθοδολογία και τα αποτελέσματα των πειραμάτων που αναπτύχθηκαν στα προηγούμενα κεφάλαια και αναλύουμε την αποτελεσματικότητα των μεθόδων και των αλγορίθμων παράγοντας και καινούρια δεδόμενα για περαιτέρω επιβεβαίωση των αποτελεσμάτων.

Τέλος στο έκτο κεφάλαιο παραθέτουμε όλους τους κώδικες που γραφτήκαν και χρησιμοποιήθηκαν στην Matlab με πλήρη σχολιασμό σε κάθε τμήμα τους και στο έβδομο κεφάλαιο βρίσκονται όλες οι διαδικτυακές και βιβλιογραφικές αναφορές.

2. Θεωρητικό υπόβαθρο

2.1. Ο χώρος $C^m(\Omega)$

Ως χώρο $C^0(\Omega)$ ορίζουμε το σύνολο :

$$C^0(\Omega) = \{u(x), \quad u(x) \text{ συνεχής στο } \Omega\}$$

, δηλαδή ο $C^0(\Omega)$ είναι το σύνολο των συναρτήσεων οι οποίες είναι συνεχείς στο Ω και είναι γραμμικός διότι από $u_1, u_2 \in C^0(\Omega)$ έχουμε ότι αu_1 και βu_2 συνεχείς, συνεπώς η $\alpha u_1 + \beta u_2$ συνεχής ή ισοδύναμα ισχύει ότι :

$$u_1, u_2 \in C^0(\Omega) \Rightarrow \alpha u_1 + \beta u_2 \in C^0(\Omega)$$

Ως $C^1(\Omega)$ ορίζουμε το σύνολο των συναρτήσεων οι οποίες είναι συνεχείς στο Ω και υπάρχει η μερική παράγωγος τους η οποία είναι επίσης συνεχής και κατ' αναλογία ορίζουμε και τους υψηλότερης τάξης χώρους όπου θα έχουμε συνεχείς παραγώγους δεύτερης τάξης, m τάξης, κ.α. δηλαδή :

$$C^1(\Omega) = \{u(x), \quad u(x) \wedge \frac{\partial u}{\partial x} \text{ συνεχής στο } \Omega\}$$

$$C^2(\Omega) = \{u(x), \quad u(x) \wedge \frac{\partial u}{\partial x} \wedge \frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}\right), \text{ συνεχής στο } \Omega\}$$

$$C^m(\Omega) = \{u(x), \quad u(x) \wedge \frac{\partial u}{\partial x} \wedge \ldots \wedge u^{(m)}(x), \text{ συνεχής στο } \Omega\}$$

, π.χ. Ο $C^3(\Omega)$ είναι ο χώρος που περιέχει τις συναρτήσεις οι οποίες είναι συνεχείς στο Ω για τις οποίες οι μερικές παράγωγοι μέχρι τρίτης τάξης υπάρχουν και είναι συνεχείς στο Ω ή ισοδύναμα οι μερικές παράγωγοι $u, u_x, u_y, u_{xx}, u_{yy}, u_{xxx}, u_{yyy}, u_{xxy}, u_{xyy}, u_{yyy}$ υπάρχουν και είναι συνεχείς στο Ω. Για μερικές παραγώγους υψηλής τάξης συναρτήσεων πολλών μεταβλητών, η σημειογραφία διευκολύνεται εάν χρησιμοποιήσουμε την εξής συνθήκη :

Έστω συνάρτηση $u(x_1,\ldots,x_n) \in R^n$ και $\alpha = (\alpha_1,\ldots,\alpha_n),\ \alpha_i \in \mathbb{Z}, \alpha_i \geq 0$, τότε :

$$D^\alpha u(x) = \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}}, \quad |\alpha| = \alpha_1 + \ldots + \alpha_n, \quad \alpha_i \geq 0$$

, π.χ. Για $n = 2$ και $u(x) = (x_1, x_2)$, οι δυνατοί συνδυασμοί είναι για το α είναι :

$$\alpha = (2,0)\,,\ \alpha = (0,2)\ ,\ \alpha = (1,1)$$

, συνεπώς για το $D^\alpha u(x)$ θα έχουμε :

$$D^{\alpha}u(x)=\frac{\partial^2 u}{\partial x_1^{\,2}}\ \left(\alpha=(2,0)\right),\quad D^{\alpha}u(x)=\frac{\partial^2 u}{\partial x_2^{\,2}}\ \left(\alpha=(0,2)\right),\quad D^{\alpha}u(x)=\frac{\partial^2 u}{\partial x_1 \partial x_2}\ \left(\alpha=(1,1)\right)$$

Εάν χρησιμοποιήσουμε την παραπάνω συνθήκη για τον χώρο $C^m(\Omega)$ θα έχουμε ότι :

$$C^m(\Omega)=\{u(x_1,\dots,x_n),\, D^{|\alpha|}u \text{ συνεχής στο } \Omega \text{ και } |\alpha|\le m\ \}$$

Περαιτέρω μπορούμε να ορίσουμε χώρους άπειρης τάξης :

$$C^{\propto}(\Omega)=\{u(x),u'(x),\dots \text{ άπειρα διαφορίσιμη στο } \Omega\}$$

, π.χ. $e^x, \sin x \in C^{\propto}(\Omega)$. Ένας διανυσματικός χώρος εφοδιασμένος με την μη αρνητική μετρική της απόστασης $d(\cdot,\cdot)$ ονομάζεται μετρικός, με την απόσταση να είναι μια συνάρτηση των διανυσμάτων u και v για την οποία ισχύει ότι :

- $d(u,v)\ge 0$
- $d(u,v)=0 \Leftrightarrow u\equiv v$
- $d(u,v+w)\le d(u,v)+d(u,w)$ (τριγωνική ανισότητα)

Ένας διανυσματικός χώρος εφοδιασμένος με την μη αρνητική συνάρτηση της νόρμας $\|\cdot\|$ ονομάζεται νορμικός, με την νόρμα να είναι μια συνάρτηση η οποία ικανοποιεί τις παρακάτω σχέσεις :

- $\|u(x)\|\ge 0,\quad \|u(x)\|=0 \Leftrightarrow u(x)\equiv 0$
- $\|\alpha\cdot u(x)\|=\alpha\|u(x)\|,\quad \alpha\in R$
- $\|u(x)+v(x)\|\le\|u(x)\|+\|v(x)\|$

Στον χώρο $C^0(\Omega)$ οι συνάρτησεις :

$$d(u,v)=\max_{x\in\Omega}|u(x)-v(x)|,\quad \|u(x)\|_{\propto}=\max_{x\in\Omega}|u(x)|$$

, είναι απόσταση και νόρμα αντίστοιχα και η επέκταση των παραπάνω συναρτήσεων στον χώρο $C^m(\Omega)$ επιτυγχάνεται με τις συναρτήσεις $d(u,v)=\max\limits_{0\le|\alpha|\le m}\max\limits_{x\in\Omega}\left|D^{\alpha}u(x)-D^{|\alpha|}v(x)\right|$ και $\|u(x)\|=\max\limits_{0\le|\alpha|\le m}\max\limits_{x\in\Omega}\left|D^{|\alpha|}u(x)\right|$ αντίστοιχα [7].

2.2. Ο χώρος $L^m(\Omega)$

Ο τετραγωνικά ολοκληρώσιμος χώρος $L^2(\Omega)$, ορίζεται ως εξής :

$$L^2(\Omega)=\left\{u(x),\quad \int_\Omega |u(x)|^2\,dx<\infty\right\}$$

, π.χ. Για την συνάρτηση $u(x)=\dfrac{1}{x}$ και $\Omega=(0,1)$:

$$\int_\Omega |u(x)|^2\,dx=\int_{0+}^{1}\frac{1}{|x|^2}dx=\lim_{t\to 0+}\int_t^1\frac{1}{x^2}dx=\lim_{t\to 0+}\int_t^1\left(-\frac{1}{x}\right)' dx=\lim_{t\to 0+}\left[-1-\left(-\frac{1}{t}\right)\right]_r^1=\lim_{t\to 0+}-1+\frac{1}{t}=+\infty$$

, συνεπώς $u(x)=\dfrac{1}{x}\notin L^2(0,1)$. Στο σημείο αυτό θέλουμε να σημειώσουμε ότι υπάρχουν συναρτήσεις οι οποίες ανήκουν στον χώρο $L^2(0,1)$ αλλά δεν ανήκουν στον χώρο $C^0(0,1)$, π.χ. Για την συνάρτηση $u(x)=\dfrac{1}{\sqrt[4]{x}}$ έχουμε ότι $u(x)=\dfrac{1}{x^{1/4}}\notin C^0(0,1)$ αλλά $\int_0^1|u(x)|^2\,dx=\int_0^1\left(\dfrac{1}{|\sqrt[4]{x}|}\right)^2 dx=\int_0^1\left(\dfrac{1}{\sqrt{x}}\right)dx=2<\infty$,

συνεπώς $u(x)\in L^2(0,1)$ και $u(x)\notin C^0(0,1)$. Στον $L^2(\Omega)$, οι συναρτήσεις :

$$d(f,g)=\left(\int_\Omega |f-g|^2\,dx\right)^{1/2},\quad \|u\|_{L^2}=\|u\|_2=\left\{\int_\Omega |u|^2\,dx\right\}^{1/2}$$

, είναι απόσταση και νόρμα αντίστοιχα [7], συνεπώς ο $L^2(\Omega)$ εφοδιασμένος με τις παραπάνω είναι μετρικός και νορμικός και βάση της απόστασης ορίζεται η ισότητα δυο συναρτήσεων στον χώρο $L^2(\Omega)$ όπου δυο συναρτήσεις f,g είναι ίσες εάν και μόνο εάν $d(f,g)=0$, π.χ. οι συναρτήσεις :

$$f(x)=\begin{cases}0, & x<0\\ 1, & x\geq 0\end{cases}\quad \text{και}\quad g(x)=\begin{cases}0, & x\leq 0\\ 1, & x>0\end{cases}$$

, είναι ίσες στον $L^2(-2,2)$, διότι :

$$d(f,g)=\left(\int_\Omega |f-g|^2\,dx\right)^{1/2}=\left(\int_{-2}^{2}|f-g|^2\,dx\right)^{1/2}+\left(\int_{-2}^{0}|f-g|^2\,dx+\int_{0}^{2}|f-g|^2\,dx\right)^{1/2}=0$$

Περαιτέρω ο χώρος $L^m(\Omega)$, ορίζεται ως εξής :

$$L^m(\Omega)=\left\{u(x),\int_\Omega |u(x)|^m\,dx<\infty\right\}$$

και η συνάρτηση $d(f,g)=\left\{\int_\Omega |f-g|^p\,dx\right\}^{1/p}$, είναι απόσταση στον χώρο $L^p(\Omega)$ [7].

Εσωτερικό γινόμενο στον χώρο $L^2(\Omega)$

Στον R^n δοθέντος δυο διανυσμάτων $x=\left[x_1,...,x_n\right]^T, y=\left[y_1,...,y_n\right]^T$, γνωρίζουμε ότι το εσωτερικό γινόμενο ορίζεται ως εξής :

$$<x,y>=x^T y=\sum_{i=1}^{n} x_i y_i$$

Κατ' αναλογία εάν ορίσουμε στον χώρο $L^2(\Omega)$ την συνάρτηση :

$$(f,g)=\int_\Omega f(x)g(x)dx, \quad f,g\in L^2(\Omega)$$

, τότε ικανοποιούνται οι συνθήκες του εσωτερικού γινομένου [7] :

- $(f,g)=(g,f), \quad \forall f,g\in L^2(\Omega)$
- $(\alpha f,g)=(f,\alpha g)=\alpha(f,g), \quad \forall f,g\in L^2(\Omega), \forall\alpha\in R$
- $(f,g+w)=(f,g)+(f,w), \quad \forall f,g\in L^2(\Omega)$

Η νόρμα, η απόσταση και το εσωτερικό γινόμενο στον χώρο $L^2(\Omega)$ συνδέονται ως εξής [7] :

$$\|u\|_2=\sqrt{(u,u)}=\left\{\int_\Omega |u|^2\,dx\right\}^{1/2}$$

2.3. Χώροι Sobolev $H^m(\Omega)$

Οι χώροι Sobolev ορίζονται χρησιμοποιώντας ως βάση τους τετραγωνικά ολοκληρώσιμα χώρους $L^2(\Omega)$ με την προσθήκη ότι πέραν της συνέχειας της συνάρτησης επιπλέον απαιτούμε οι μερικές παράγωγοι της συνάρτησης να ανήκουν στον $L^2(\Omega)$, δηλαδή κατασκευάζουμε χώρους συναρτήσεων με παραγώγους σε μορφή ολοκληρωμάτων.

Ορισμός χώρων Sobolev

Ως μηδενικής τάξης χώρο Sobolev $H^0(\Omega)$, ορίζουμε το σύνολο των συναρτήσεων οι οποίες ανήκουν στον $L^2(\Omega)$ για τις οποίες δεν υπάρχουν οι μερικές παράγωγοι (τετριμένη περίπτωση), συνεπώς προκύπτει ότι είναι της μορφής :

$$H^0(\Omega)=L^2(\Omega)=\left\{u(x), \quad \int_\Omega |u(x)|^2 <\infty\right\}$$

Ως πρώτης τάξης χώρο Sobolev $H^1(\Omega)$, ορίζουμε το σύνολο των συναρτήσεων οι οποίες ανήκουν στον $L^2(\Omega)$ και υπάρχουν μερικές παραγώγοι πρώτης τάξης οι οποίες ανήκουν στον $L^2(\Omega)$, συνεπώς προκύπτει ότι είναι της μορφής :

$$H^1(\Omega) = \left\{ v(x),\ D^{|\alpha|}v \in L^2(\Omega), |\alpha| \le 1 \right\}$$

, π.χ. για $\Omega = (\alpha, b)$ προκύπτει ότι ο χώρος $H^1(\alpha, b)$ ορίζεται ως εξής :

$$H^1(\alpha,b) = \left\{ v(x),\ \alpha < x < b \ \wedge\ \int_\alpha^b |v(x)|^2\, dx < \propto \ \wedge\ \int_\alpha^b \left|\frac{\partial v}{\partial x}\right|^2 dx < \propto \right\}$$

, με την συνάρτηση $f(x) = x^2$ να ανήκει στον $H^1(\alpha, b)$ διότι :

$$\int_0^1 |x|^2 dx = \frac{1}{3} < \propto \ \wedge\ \int_0^1 \left|\frac{\partial f}{\partial x}\right|^2 dx = \int_0^1 (2x)^2 dx = 4\int_0^1 x^2 dx = \frac{4}{3} < \propto$$

Η επέκταση του παραπάνω ορισμού στις δυο διαστάσεις σχηματίζεται ως εξής :

$$H^1(\Omega) = \left\{ v(x,y),\ \ \alpha < x < b \wedge v(x,y) \in L^2(\Omega) \wedge \frac{\partial v}{\partial x} \in L^2(\Omega) \wedge \frac{\partial v}{\partial y} \in L^2(\Omega) \right\}$$

, π.χ. Για την συνάρτηση $g(x,y) = \eta\mu(x_1) + \sigma\upsilon\nu(x_2)$ και $\Omega = [0, 2\pi]\ x\ [0, 2\pi]$, ισχύει ότι :

$$\int_\Omega |g(x,y)|^2\, dx = \int_\Omega |\eta\mu(x_1) + \sigma\upsilon\nu(x_2)|^2\, dx \le \int_\Omega |1+1|^2\, dx < \propto \ \Rightarrow\ g(x,y) \in L^2(\Omega)$$

$$\int_\Omega \left|\frac{\partial g}{\partial x_1}\right|^2 dx = \int_\Omega |\sigma\upsilon\nu(x_1)|^2\, dx \le \int_\Omega |1|^2\, dx < \propto \ \Rightarrow\ \frac{\partial g}{\partial x_1} \in L^2(\Omega)$$

$$\int_\Omega \left|\frac{\partial g}{\partial x_2}\right|^2 dx = \int_\Omega |-\eta\mu(x_2)|^2\, dx \le \int_\Omega |1|^2\, dx < \propto \ \Rightarrow\ \frac{\partial g}{\partial x_2} \in L^2(\Omega)$$

, συνεπώς $g(x,y) \in H^1(\Omega)$. Εν γένει ως m τάξης χώρο Sobolev $H^m(\Omega)$ ορίζουμε το σύνολο των συναρτήσεων οι οποίες ανήκουν στον $L^2(\Omega)$ και υπάρχουν μερικές παραγώγοι μέχρι m τάξη οι οποίες ανήκουν στον $L^2(\Omega)$, συνεπώς προκύπτει ότι είναι της μορφής :

$$H^m(\Omega) = \left\{ v(x),\ D^{|\alpha|}v \in L^2(\Omega), |\alpha| \le m \right\}$$

Εσωτερικό γινόμενο, απόσταση και νόρμα στους χώρους Sobolev

Εφόσον ο χώρος $H^0(\Omega)$ ταυτίζεται με τον $L^2(\Omega)$ όπου το εσωτερικό γινόμενο είναι $(f,g)=\int_\Omega f(x)g(x)dx$, για τον $H^0(\Omega)$ προκύπτει ότι :

$$(u,v)_{H^0(\Omega)}=(u,v)_{L^2(\Omega)}=\int_\Omega uvdx$$

Το εσωτερικό γινόμενο στον χώρο $H^1(\Omega)$ ορίζεται επεκτείνοντας τον παραπάνω ορισμό με την πρόσθεση του γινουμένου της πρώτης τάξης παραγώγων, π.χ. εάν θέσουμε $\Omega=(\alpha,b)$, τότε το εσωτερικό γινόμενο των συναρτήσεων u και v στον χώρο $H^1(\alpha,b)$ είναι $(u,v)_{H^1(\alpha,b)}=\int_\alpha^b (uv+u'v')dx$ και η επέκταση του στις δυο διαστάσεις, ορίζεται κατ' αναλογία ως εξής :

$$(u,v)_{H^1(\Omega)}=\iint_\Omega\left(uv+\frac{\partial u\partial v}{\partial x\partial x}+\frac{\partial u\partial v}{\partial y\partial y}\right)dxdy$$

Εν γένει στον χώρο $H^m(\Omega)$ το εσωτερικό γινόμενο κατασκευάζεται με την πρόσθεση των όρων του αθροίσματος που ολοκληρώνουμε επί του πεδίου Ω μέχρι και τις τάξης m μερικές παραγώγους ή ισοδύναμα είναι της μορφής :

$$(u,v)_{H^m(\Omega)}=(u,v)_m=\iint_\Omega\sum_{|\alpha|\le m}\left(D^{|\alpha|}u(x)\right)\left(D^{|\alpha|}v(x)\right)dx$$

Για την νόρμα και την απόσταση στον χώρο Sobolev $H^m(\Omega)$, οι παρακάτω συναρτήσης εκπληρούν τις προυποθέσεις που προαναφέραμε [8] :

$$\|u\|_{H^m(\Omega)}=\|u\|_m=\left\{\int_\Omega\sum_{|\alpha|\le m}\left|D^{|\alpha|}u(x)\right|^2dx\right\}^{1/2},\ d(u,v)_m=\|u-v\|_m$$

, π.χ. Εάν θέσουμε $\Omega=(\alpha,b)$ στον χώρο $H^1(\alpha,b)$ δυο συναρτήσεις που είναι νόρμα και απόσταση αντίστοιχα είναι οι εξής [8] :

$$\|u\|_1=\left\{\int_\alpha^b\left[(u^2)+(u')^2\right]dx\right\}^{1/2},\ d(u,v)_1=\|u-v\|_1=\left\{\int_\alpha^b\left[(u-v)^2+((u-v)')^2\right]dx\right\}^{1/2}$$

2.4. Ασθενής παράγωγος και το θεώρημα ενσωμάτωσης του Sobolev

Ορίζουμε ως ασθενή παράγωγο (weak derivative) m τάξης της $u(x) \in H^0(\Omega)$ και συμβολίζουμε με $v(x) = u^{(m)}(x)$, μια συνάρτηση $v(x)$ η οποία ικανοποιεί την συνθήκη :

$$\int_\Omega v(x)\varphi(x)dx = (-1)^m \int_\Omega u(x)\varphi^{(m)}(x)dx$$

, όπου για την $\varphi(x)$ ισχύει ότι :

$$\varphi(x) \in C^m(\Omega) \quad \wedge \quad \varphi(x) = \varphi'(x) = ... = \varphi^{(m-1)}(x) = 0, \; x \in \partial\Omega$$

, με $\partial\Omega$ να είναι το σύνορο του Ω και την $\varphi(x)$ να συναντάται στην βιβλιογραφία με τον όρο συνάρτηση δοκιμής (testing function), π.χ. για $\Omega = (0,1)$ και m=1 η ασθενής παράγωγος πρώτης τάξης $v(x) = u^{(1)}(x)$ θα είναι μια συνάρτηση $v(x)$, τέτοια ώστε :

$$\int_\Omega v(x)\varphi(x)dx = -\int_0^1 u^{(1)}(x)\varphi(x)dx$$

, όπου για την συνάρτηση δοκιμής $\varphi(x)$ θα ισχύει ότι :

$$\varphi(x) \in C^1(0,1) \quad \wedge \quad \varphi(0) = \varphi(1) = \varphi'(0) = \varphi'(1) = 0$$

Η κατα παράγοντες ολοκλήρωση ισχύει και για τις ασθενείς παραγώγους και παρατηρούμε ότι ο ορισμός προκύπτει ως εξής [19] :

$$\int_\Omega v(x)\varphi(x)dx = \int_0^1 u^{(1)}(x)\varphi(x)dx = \int_0^1 u'(x)\varphi(x)dx = [u\varphi]_0^1 - \int_0^1 u(x)\varphi'(x)dx \Rightarrow$$
$$\int_\Omega v(x)\varphi(x)dx = -\int_0^1 u(x)\varphi'(x)dx = (-1)^{(1)} \int_\Omega u(x)\varphi^{(1)}(x)dx$$

, π.χ. Η συνάρτηση

$$u(x) = \begin{cases} 0, & x \le 0 \\ x, & x > 0 \end{cases}$$

, έχει ως ασθενή παράγωγο την συνάρτηση :

$$v(x) = \begin{cases} 0, & x \le 0 \\ 1, & x > 0 \end{cases}$$

Για να αποδείξουμε ότι η v(x) είναι η ασθενής παράγωγος της u(x), πολλαπλασιάζουμε την v με μια δοκιμαστική συνάρτηση φ και ολοκληρώνουμε επί του πεδίου ορισμού :

$$\int_R v(x)\varphi(x)dx = \int_0^{\infty}\varphi(x)dx = \int_0^{\infty} x'\varphi(x)dx = \lim_{t\to\infty}[x\varphi(x)]_0^t - \int_0^{\infty} x\varphi'(x)dx \Rightarrow$$

$$\int_R v(x)\varphi(x)dx = -\int_0^{\infty} u(x)\varphi'(x)dx = \left(-1\right)^{(1)}\int_{\Omega} u(x)\varphi^{(1)}(x)dx$$

, διότι η συνάρτηση δοκιμής στην συνοριακή περιοχή είναι μηδενική δηλαδή ισχύει ότι :

$$\lim_{t\to\infty}\varphi(x) = \lim_{t\to\infty}\varphi'(x) = ... = 0$$
$$\lim_{t\to-\infty}\varphi(x) = \lim_{t\to-\infty}\varphi'(x) = ... = 0$$

, συνεπώς η v(x) είναι ασθενής παράγωγος της u(x) για κάθε δοκιμαστική συνάρτηση η οποία πληρεί τις παραπάνω συνθήκες. Παρακάτω παραθέτουμε το θεώρημα ενσωμάτωσης του Sobolev, το οποίο συνδέει τους χώρους $H^m(\Omega)$ οι οποίοι ορίστηκαν βάσει των χώρων $L^2(\Omega)$ με τους χώρους $C^m(\Omega)$.

Θεώρημα ενσωμάτωσης Sobolev [9]. Εάν $2m > n$, τότε :

$$H^{m+j}(\Omega) \subset C^j(\Omega), \quad j = 0,1,...$$

, με n να είναι η διάσταση της ανεξάρτητης μεταβλητής των στοιχείων του χώρου Sobolev.

• Σε χώρους μιας διάστασης έχουμε ότι $n = 1$ συνεπώς $m > \frac{1}{2} \Rightarrow m \geq 1$ και από το θεώρημα ενσωμάτωσης προκύπτει ότι :

$$H^{1+j}(\Omega) \overset{j=0}{=} H^{1+0}(\Omega) = H^1(\Omega) \subset C^0(\Omega)$$
$$H^{1+j}(\Omega) \overset{j=1}{=} H^{1+1}(\Omega) = H^2(\Omega) \subset C^1(\Omega)$$
$$H^{1+j}(\Omega) \overset{j=2}{=} H^{1+2}(\Omega) = H^3(\Omega) \subset C^2(\Omega)$$
$$\vdots$$

Να σημειώσουμε ότι εάν θέταμε ως $m = 2$, τότε προκύπτει :

$$H^{2+j}(\Omega) \overset{j=0}{=} H^{2+0}(\Omega) = H^2(\Omega) \subset C^0(\Omega)$$
$$H^{2+j}(\Omega) \overset{j=1}{=} H^{2+1}(\Omega) = H^3(\Omega) \subset C^1(\Omega)$$
$$\vdots$$

, το οποίο προκύπτει και από το γεγονός $C^1(\Omega) \subset C^0(\Omega)$.

• Σε χώρους δυο διαστάσεων έχουμε ότι $n=2$ συνεπώς $m>1 \Rightarrow m \geq 2$ και από το θεώρημα ενσωμάτωσης προκύπτει ότι :

$$\begin{aligned}
&H^{2+j}(\Omega) \overset{j=0}{=} H^{2}(\Omega) = H^{2+0}(\Omega) \subset C^{0}(\Omega) \\
&H^{2+j}(\Omega) \overset{j=1}{=} H^{3}(\Omega) = H^{2+1}(\Omega) \subset C^{1}(\Omega) \\
&H^{2+j}(\Omega) \overset{j=2}{=} H^{4}(\Omega) = H^{2+2}(\Omega) \subset C^{2}(\Omega) \\
&\qquad\qquad \vdots
\end{aligned}$$

• Στον χώρο των τριών διαστάσεων έχουμε ότι $n=3$ συνεπώς $m>\frac{3}{2} \Rightarrow m \geq 2$ και από το θεώρημα ενσωμάτωσης προκύπτει ότι :

$$\begin{aligned}
&H^{2+j}(\Omega) \overset{j=0}{=} H^{2}(\Omega) = H^{2+0}(\Omega) \subset C^{0}(\Omega) \\
&H^{2+j}(\Omega) \overset{j=1}{=} H^{3}(\Omega) = H^{2+1}(\Omega) \subset C^{1}(\Omega) \\
&H^{2+j}(\Omega) \overset{j=2}{=} H^{4}(\Omega) = H^{2+2}(\Omega) \subset C^{2}(\Omega) \\
&\qquad\qquad \vdots
\end{aligned}$$

2.5. Συμμετρικό και ασύμμετρο πρόβλημα μεταβολών

Διγραμμικές μορφές

Σε έναν διανυσματικό χώρο V και ένα πεδίο F ως διγραμμική μορφή ορίζουμε μια συνάρτηση $\alpha : VxV \to F$, η οποία είναι γραμμική ως προς κάθε μεταβλητή, δηλαδή ισχύει ότι :

$$\begin{aligned}
&\alpha(u, v+w) = \alpha(u,v) + \alpha(u,w) \\
&\alpha(u+v, w) = \alpha(u,w) + \alpha(v,w) \\
&\alpha(\lambda u, v) = \alpha(u, \lambda v) = \lambda \alpha(u,v)
\end{aligned}$$

, π.χ. Η συνάρτηση $\alpha : RxR \to R$ με κανόνα αντιστοίχησης $\alpha(x,y) = \lambda xy,\ \lambda \in R$ είναι διγραμμική μορφή, διότι :

$$\begin{aligned}
&\alpha(u, v+w) = \lambda u(v+w) = \lambda uv + \lambda uw = \alpha(u,v) + \alpha(u,w) \\
&\alpha(u+v, w) = \lambda(u+v)w = \lambda uw + \lambda vw = \alpha(u,w) + \alpha(v,w) \\
&\alpha(ku, v) = \lambda\left(kuv\right) = \lambda k\left(uv\right) \\
&\alpha(u, \lambda v) = \lambda u\left(kv\right) = \lambda k\left(uv\right) \\
&\lambda\alpha(u,v) = \lambda\left(kuv\right) = \lambda k\left(uv\right)
\end{aligned}$$

• Μια διγραμμική μορφή $\alpha(\cdot,\cdot)$ είναι **συμμετρική** εάν :

$$\alpha(u,v)=\alpha(v,u) \quad \forall u,v\in V$$

, π.χ. Η συνάρτηση $\alpha: RxR\to R$ με κανόνα αντιστοίχησης $\alpha(x,y)=\lambda xy,\ \lambda\in R$ είναι συμμετρική διγραμμική μορφή, διότι :

$$\alpha(x,y)=\lambda xy=\lambda yx=\alpha(y,x),\ \lambda\in R$$

Σε κάθε συμμετρική διγραμμική μορφή $\alpha(\cdot,\cdot)$ αντιστοιχεί μια τετραγωνική μορφή $Q(x)$ με την εξής σχέση [10] :

$$\begin{aligned}&Q(x)=\alpha(x,x)\\&\alpha(x,y)=\alpha(y,x)=\tfrac{1}{2}(Q(x+y)-Q(x)-Q(y))\end{aligned}$$

• Μια διγραμμική μορφή $\alpha(\cdot,\cdot)$ είναι **θετική** εάν :

$$\alpha(u,u)\geq 0, \quad \forall u\in V$$

, π.χ. Η συνάρτηση $\alpha: RxR\to R$ με κανόνα αντιστοίχησης $\alpha(x,y)=xy$ είναι διγραμμική και θετική μορφή, διότι :

$$\alpha(x,x)=x^2\geq 0$$

• Μια διγραμμική μορφή $\alpha(\cdot,\cdot)$ είναι **φραγμένη** (ή συνεχής) εάν υπάρχει σταθερά $C>0$ τέτοια ώστε :

$$\alpha(u,v)\leq C\|u\|\|v\|, \quad \forall u,v\in V$$

, π.χ. Η συνάρτηση $\alpha: R^+xR^+\to R^+$ με κανόνα αντιστοίχησης $\alpha(x,y)=xy$ είναι διγραμμική και για $u,v\in R^*$, θα έχουμε ότι :

$$\begin{aligned}&\alpha(u,v)=xy\\&\|u\|\|v\|=x^2y^2=\left(xy\right)^2=\left(\alpha(u,v)\right)^2\end{aligned}$$

, συνεπώς θα έχουμε ότι $\alpha(u,v)\leq C\|u\|\|v\|, \quad C=\dfrac{1}{\alpha(u,v)}=\dfrac{1}{xy}>0$.

• Μια διγραμμική μορφή $\alpha(u,u)$ είναι **k – ελλεπτική** εάν υπάρχει σταθερά $k > 0$, τέτοια ώστε:

$$\alpha(u,u) \geq k\|u\|^2, \quad \forall u \in V$$

, π.χ. Η συνάρτηση $\alpha : RxR \to R$ με κανόνα αντιστοίχησης $\alpha(x,y) = \kappa(xy), \ \kappa > 0$ είναι διγραμμική και για $u = x \in R$, θα έχουμε ότι:

$$\alpha(u,u) = \kappa x^2$$
$$\|u\|^2 = x^2$$

, συνεπώς θα έχουμε ότι $\alpha(u,v) = \kappa x^2 \geq \kappa\|u\|^2, \quad \kappa > 0$.

Συμμετρικό και ασύμμετρο πρόβλημα μεταβολών

Πρίν παρουσιάσουμε τις διατυπώσεις των σχετικών προβλήματων, παραθέτουμε την παρακάτω πρόταση η οποία χρησιμοποιείται και στις δυο αποδείξεις που αναφέρονται στην σχετική βιβλιογραφία.

Πρόταση Α [11]. Έστω ο χώρος Hilbert H εφοδιασμένος με μια νόρμα $\|\cdot\|$**,** εσωτερικό γινόμενο $\langle \ \rangle$ και $\alpha : HxH \to F$ μια διγραμμική μορφή στο $V \subset H$, τέτοια ώστε να είναι φραγμένη και k – ελλεπτική. Ο χώρος V εφοδιασμένος με την διγραμμική μορφή $(V, \alpha(\cdot,\cdot))$ είναι χώρος Hilbert.

Συμμετρικό πρόβλημα μεταβολών [11]

Εάν έχουμε ότι :

1. Ο χώρος $(H, (\cdot,\cdot))$ είναι ενας χώρος Hilbert στο R.
2. V είναι ενα κλειστό υποσύνολο του H.
3. Η διγραμμική μορφή $\alpha : VxV \to R$ είναι φραγμένη.
4. Η διγραμμική μορφή $\alpha : VxV \to R$ είναι k - ελλειπτική στο V.
5. Η διγραμμική μορφή $\alpha : VxV \to R$ είναι συμμετρική.

, τότε δοθέντος $f \in V'$ να βρεθεί $u \in V$ τέτοιο ώστε:

$$\alpha(u,v) = f(v), \quad \forall v \in V$$

, όπου V' είναι ο χώρος των συνεχών γραμμικών συναρτήσεων από το V στο R. Το συμμετρικό προβλήμα έχει μοναδική λύση [11] και είναι η βάση για την ανάπτυξη της μεθόδου Ritz, όπου ο χώρος V' στην βιβλιογραφία συναντάται ως V* όταν αναφερόμαστε γενικά σε πεδίο F και ως V' όταν αναφερόμαστε σε κάποιο συγκεκριμένο πεδίο. Σημείωνουμε για την έννοια της συνέχειας στην τοπολογία, ότι μια συνάρτηση $f : X \to Y$ ορίζεται ως συνεχής εάν για κάθε ανοικτό σύνολο $V \subseteq Y$ η αντίστροφη εικόνα $f^{-1}(V) = \{x \in X, f(x) \in V\}$, είναι ανοιχτό υποσύνολο του X, όπου για την f στο συγκεκριμένο πρόβλημα έχουμε $f : V \to R$.

Ασύμμετρο πρόβλημα μεταβολών [11]

Εάν έχουμε ότι:

1. Ο χώρος $(H,(\cdot,\cdot))$ είναι ενας χώρος Hilbert στο R.

2. V είναι ενα κλειστό υποσύνολο του H

3. Η διγραμμική μορφή $\alpha : VxV \to R$ είναι φραγμένη

4. Η διγραμμική μορφή $\alpha : VxV \to R$ είναι k - ελλειπτική στο V

, τότε δοθέντος $f \in V'$ να βρεθεί $u \in V$ τέτοιο ώστε:

$$\alpha(u,v) = f(v), \quad \forall v \in V$$

, όπου V' είναι ο χώρος των συνεχών γραμμικών συναρτήσεων.

Το ασύμμετρο προβλήμα είναι η βάση για την ανάπτυξη της μεθόδου Galerkin και έχει μοναδική λύση λόγω του παρακάτω θεωρήματος :

Θεώρημα Lax - Milgram [12]. Έστω ο χώρος Hilbert H εφοδιασμένος με νόρμα $\|\cdot\|$**,** εσωτερικό γινόμενο $\langle\ \rangle$ και $\alpha : HxH \to R$ μια διγραμμική μορφή τέτοια ώστε :

1. Να είναι φραγμένη, δηλαδή υπάρχει σταθερά $C > 0$ τέτοια ώστε: $\alpha(u,v) \le C\|u\|\|v\|, \quad \forall u,v \in V$

2. Να είναι k – ελλειπτική, δηλαδή υπάρχει σταθερά $k > 0$, τέτοια ώστε : $k\|u\|^2 \le \alpha(u,u)$.

Εάν έχουμε μια γραμμική συνάρτηση $1 : H \to R$, τότε υπάρχει μοναδική λύση u στην εξίσωση:

$$\alpha(u,v) = l(v), \quad \forall v \in H$$

2.6. Μέθοδος Ritz

Με την μέθοδο Ritz επιδιώκουμε να λύσουμε το συμμέτρικο πρόβλημα μεταβολών, δηλαδή αναζητούμε $u_h \in V_h$ για την οποία ισχύει $\alpha(u_h, v_h) = l(v_h), \quad \forall v_h$ και έχει ως βάση την παρακάτω πρόταση.

Πρόταση Β [12]. Εάν έχουμε μια συμμετρική και θετική διγραμμική μορφή, τότε η εξίσωση :

$$\alpha(u,v) = l(v), \quad \forall v \in H$$

έχει λύση u εάν και μόνο εάν η u ελαχιστοποιεί στο V την παράσταση :

$$J(u) = \inf_{vEV} J(v)$$

, όπου η $J(v)$ ορίζεται ως εξής :

$$J(v) = \frac{1}{2}\alpha(v,v) - l(v), \quad \forall v \in V$$

Εάν συνδυάσουμε την παραπάνω πρόταση με τις συνθήκες και την μοναδικότητα της λύσης του συμμετρικού προβλήματος, προκύπτει ότι :

Πρόταση Γ [12]. Εάν έχουμε μια φραγμένη, k - ελλειπτική, συμμετρική και θετική διγραμμική μορφή, τότε η εξίσωση:

$$\alpha(u, v) = l(v), \quad \forall v \in H$$

έχει λύση u η οποία είναι μοναδική εάν και μόνο εάν η u ελαχιστοποιεί την παράσταση :

$$J(u) = \inf_{vEV} J(v)$$

, με $J(v) = \frac{1}{2}\alpha(v, v) - l(v), \quad \forall v \in V$.

Η μέθοδος Ritz έχει ως πυρήνα την αντικατάσταση του χώρου V, με έναν πεπερασμένων διαστάσεων χώρο V_h και την εύρεση μιας προσέγγιστικής λύσης $u_h \in V_h$, τέτοιας ώστε :

$$J(u_h) = \inf_{vEV} J(v_h), \; J(v_h) = \frac{1}{2}\alpha(v_h, v_h) - l(v_h)$$

Για την εύρεση της αριθμητικής τιμής της λύσης η κεντρική ιδέα της μεθόδου είναι να ανάγουμε το πρόβλημα από έναν χώρο V σε έναν χώρο V_h ο οποίος όμως είναι πεπερασμένων διαστάσεων, συνεπώς μπορούμε να υπολογίσουμε μια βάση του και κατ' επέκταση να εκφράσουμε κάθε στοιχείο του χώρου ως γραμμικό συνδυασμό αυτής. Έστω ότι η διάσταση του V_h είναι N και ότι τα διανύσματα βάσης είναι της μορφής $\varphi_j, \; j = 1, ..., N$, τότε θα έχουμε ότι κάθε στοιχείο του χώρου είναι της μορφής $u_h = \sum_{j=1}^{N} u_j\varphi_j, \quad ,\forall u_h \in V_h$ και για την λύση u_h που αναζητούμε, θα πρέπει :

$$J(u_h) = \inf_{vEV} J(v_h), \; J(v_h) = \frac{1}{2}\alpha(v_h, v_h) - l(v_h)$$

, συνεπώς αντικαθιστώντας λαμβάνουμε ότι :

$$J(u_h) = \frac{1}{2}\alpha\left(\sum_{j=1}^{N} u_j\varphi_j, \sum_{j=1}^{N} u_j\varphi_j\right) - l\left(\sum_{j=1}^{N} u_j\varphi_j\right) = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} u_i u_j \alpha\left(\varphi_i, \varphi_j\right) - \sum_{j=1}^{N} u_j l\left(\varphi_j\right)$$

Η παραπάνω ισότητα προκύπτει διότι αφενός διότι η l είναι γραμμική μορφή, συνεπώς :

$$l\left(\sum_{j=1}^{N} u_j\varphi_j\right) = l\left(u_1\varphi_1 + ... + u_N\varphi_N\right) = l\left(u_1\varphi_1\right) + ... + l\left(u_N\varphi_N\right) = u_1 l\left(\varphi_1\right) + ... + u_N l\left(\varphi_N\right) = \sum_{j=1}^{N} u_j l\left(\varphi_j\right)$$

, αφετέρου εάν για την ανάλυση της παράστασης :

$$\frac{1}{2}\alpha\left(\sum_{j=1}^{N}u_j\varphi_j,\sum_{j=1}^{N}u_j\varphi_j\right)$$

, χρησιμοποιούμε την γραμμικότητα ως προς και τις δυο μεταβλητές :

$$\frac{1}{2}\alpha\left(\sum_{j=1}^{N}u_j\varphi_j,\sum_{j=1}^{N}u_j\varphi_j\right)=\frac{1}{2}\alpha\left(u_1\varphi_1+...+u_N\varphi_N,u_1\varphi_1+...+u_N\varphi_N\right)\Leftrightarrow$$

$$\frac{1}{2}\alpha\left(\sum_{j=1}^{N}u_j\varphi_j,\sum_{j=1}^{N}u_j\varphi_j\right)=\frac{1}{2}\left(\alpha\left(u_1\varphi_1+...+u_N\varphi_N,u_1\varphi_1\right)+...+\alpha\left(u_1\varphi_1+...+u_N\varphi_N,u_N\varphi_N\right)\right)\Leftrightarrow$$

$$\frac{1}{2}\alpha\left(\sum_{j=1}^{N}u_j\varphi_j,\sum_{j=1}^{N}u_j\varphi_j\right)=\frac{1}{2}\left(\alpha\left(u_1\varphi_1,u_1\varphi_1\right)+...+\alpha\left(u_N\varphi_N,u_1\varphi_1\right)+...+\alpha\left(u_1\varphi_1,u_N\varphi_N\right)+...+\alpha\left(u_N\varphi_N,u_N\varphi_N\right)\right)\Leftrightarrow$$

$$\frac{1}{2}\alpha\left(\sum_{j=1}^{N}u_j\varphi_j,\sum_{j=1}^{N}u_j\varphi_j\right)=\frac{1}{2}\left(u_1u_1\alpha\left(\varphi_1,\varphi_1\right)+...+u_1u_N\alpha\left(\varphi_N,\varphi_1\right)+...+u_Nu_1\alpha\left(\varphi_1,\varphi_N\right)+...+u_Nu_N\alpha\left(\varphi_N,\varphi_N\right)\right)\Leftrightarrow$$

$$\frac{1}{2}\alpha\left(\sum_{j=1}^{N}u_j\varphi_j,\sum_{j=1}^{N}u_j\varphi_j\right)=\frac{1}{2}\left(u_1u_1\alpha\left(\varphi_1,\varphi_1\right)+...+u_1u_N\alpha\left(\varphi_1,\varphi_N\right)+...+u_Nu_1\alpha\left(\varphi_N,\varphi_1\right)+...+u_Nu_N\alpha\left(\varphi_N,\varphi_N\right)\right)\Leftrightarrow$$

$$\frac{1}{2}\alpha\left(\sum_{j=1}^{N}u_j\varphi_j,\sum_{j=1}^{N}u_j\varphi_j\right)=\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}u_iu_j\alpha\left(\varphi_i,\varphi_j\right)$$

Εάν θέσουμε:

$$A=(\alpha_{ij})\in R^{N,N},\ \ \alpha_{ij}=\alpha\left(\varphi_i,\varphi_j\right)$$
$$F=(f_i)\in R^{N},\ \ f_i=l\left(\varphi_i\right)$$

, τότε το σύστημα γράφεται σε μορφή πινάκων $AU=F$ και για το πρόβλημα ελαχιστοποίησης της παράστασης $J(u_h)=\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}u_iu_j\alpha\left(\varphi_i,\varphi_j\right)-\sum_{j=1}^{N}u_jl\left(\varphi_j\right)$ έχουμε ότι ανάγεται στην εύρεση $U=(u_1,...,u_N)\in R^N$, τέτοιου ώστε :

$$J(U)=\inf_{V\in R^N}J(V)$$

, όπου για την $J(V)$ ισχύει ότι:

$$J(V)=\frac{1}{2}V^tAV-VF$$

Η υπολογιστική αντιμετώπιση του προβλήματος βασίζεται στο γεγονός ότι όπως προαναφέραμε κάθε διγραμμική μορφή αντιστοιχίζεται σε μια τετραγωνική μορφή και εφόσον η διγραμμική μορφή είναι θετικά ορισμένη και συμμετρική ο πίνακας A με $A=(\alpha_{ij})\in R^{N,N},\ \alpha_{ij}=\alpha\left(\varphi_i,\varphi_j\right)$**,** θα είναι θετικά ορισμένος και συμμετρικός. Από τα προθύστερα προκύπτει ότι για να λύσουμε το σύστημα $AU=F$ πρέπει να ελαχιστοποιήσουμε την τετραγωνική μορφή $\frac{1}{2}V^tAV-VF$, όμως ο πίνακας A είναι θετικά ορισμένος και συμμετρικός συνεπώς γνωρίζουμε ότι η λύση του προβλήματος ελαχιστοποίησης $U=(u_1,...,u_N)\in R^N$ υπάρχει, είναι μοναδική και είναι η λύση του συστήματος $AU=F$ [13].

Συνοψίζοντας η μέθοδος Ritz μετατοπίζει το πρόβλημα επίλυσης της $\alpha(u,v)=l(v)$ σε έναν πεπερασμένο χώρο και το ανάγει στην λύση του συστήματος $AU=F$, το οποίο όμως αντιστοιχεί στην ελαχιστοποίηση μιας τετραγωνικής μορφής με θετικά ορισμένο και συμμετρικό πίνακα A, συνεπώς θα υπάρχει μοναδική λύση.

2.7. Μέθοδος Galerkin

Η μέθοδος Galerkin βασίζεται στην κατασκευή του πεπερασμένου χώρου V_h, εντός του οποίου θα αναζητήσουμε την προσεγγιστική λύση και για την εύρεση της αριθμητικής τιμής της λύσης η κεντρική ιδέα της μεθόδου ταυτίζεται με αυτήν της μεθόδου Ritz, δηλαδή ανάγουμε το πρόβλημα από έναν χώρο V σε έναν χώρο V_h ο οποίος όμως είναι πεπερασμένων διαστάσεων, συνεπώς μπορούμε να υπολογίσουμε μια βάση του και κατ' επέκταση να εκφράσουμε κάθε στοιχείο του χώρου ως γραμμικό συνδυασμό αυτής [12]. Με την μέθοδο Galerkin επιδιώκουμε να λύσουμε το ασσύμετρο πρόβλημα μεταβολής problem, δηλαδή αναζητούμε $u_h\in V_h$ για την οποία ισχύει :

$$\alpha(u_h,v_h)=l(v_h),\quad \forall v_h$$

Εφόσον οι χώροι είναι πεπερασμένοι μπορούμε να θεωρήσουμε τις συναρτήσεις βάσης του κάθε χώρου $\varphi_j,\ \ j=1,...,N \ \wedge\ \psi_j,\ \ j=1,...,N$ αντίστοιχα, συνεπώς το παραπάνω πρόβλημα ανάγεται στην εύρεση $u_h=\sum_{j=1}^{N}u_j\varphi_j\in V_h$, τέτοιο ώστε :

$$\sum_{j=1}^{N}u_j\alpha(\varphi_j,\psi_i)=l(\psi_i),\quad i=1,...,N$$

ή εάν θέσουμε $A=(\alpha_{ij})\in R^{N,N},\ \alpha_{ij}=\alpha\left(\varphi_i,\varphi_j\right)$ και $F=(f_i)\in R^N,\ \ f_i=l\left(\varphi_i\right),\ i=1,...,N$, το πρόβλημα ανάγεται στο να λύσουμε το σύστημα $AU=F$. Η μεθόδος στηρίζεται στο γεγονός στο ότι τα στοιχεία του πίνακα A αντιστοιχίζονται σε διγραμμική μορφή, συνεπώς θα έχουμε ύπαρξη και μοναδικότητα της λύσης λόγω του θεωρήματος Lax – Milgram.

Συνοψίζοντας η μέθοδος Galerkin μετατοπίζει το πρόβλημα επίλυσης της $\alpha(u,v)=l(v)$ σε έναν πεπερασμένο χώρο και το ανάγει στην λύση του συστήματος $AU=F$ με τα στοιχεία του πίνακα να βρίσκονται σε πλήρη αντιστοιχία με μια διγγραμική και μια γραμμική μορφή, συνεπώς έχουμε ύπαρξη και μοναδικότητα της λύσης λόγω του θεωρήματος Lax – Milgram.

2.8. Η μέθοδος των πεπερασμένων στοιχείων

Η μέθοδος των πεπερασμένων στοιχείων μπορεί να διασπαστεί σε τέσσερις λογικές ενότητες:

Ενότητα A – Μέθοδοι

Εφαρμόζουμε την μέθοδο Ritz ή την μέθοδο Galerkin, συνεπώς το αρχικό πρόβλημα επίλυσης εξίσωσης με μερικές παραγώγους ανάγεται στην λύση του συστήματος $AU = F$ σε έναν πεπερασμένο χώρο.

Ενότητα B – Μεθοδολογία λύσης του συστήματος

1ο Βήμα: Τμηματοποιούμε το πεδίο ορισμού Ω με ενα πλέγμα T_h, δηλαδή με μια συλλογή από απλού τύπου στοιχεία με την παράμετρο h να ορίζεται ως εξής :

$$H = \max_{K \in T_h} \left(\mathrm{diam}(K) \right)$$

, συνεπώς η H μας δίνει την εκλέπτυνση του πλέγματος και όταν $H \to 0$, ο V_h θα τείνει στον V.

2ο Βήμα: Σε κάθε ενα στοιχείο υπολογίζουμε τις συναρτήσεις βάσης όπως αναλύουμε στην παρακάτω ενότητα και εφόσον κάθε σημείο εντός του στοιχείου μπορεί να εκφραστεί ως γραμμικός συνδυασμός τους, η λύση $u_h \in V_h$ που περιέχει όλες τις συναρτήσεις βάσης θα είναι η προσέγγιση μας.

Ενότητα Γ – Υπολογισμός των συναρτήσεων βάσης

1. Κατασκευάζουμε ενα πλέγμα για την υπο εκτίμηση συνάρτηση και επιλέγουμε κάποια μορφή για τα πεπερασμένα στοιχεία, π.χ. Τριγωνικά, κυβικά. κ.α.
2. Για να υπολογίσουμε τις συναρτήσεις βάσης κάθε στοιχείου ακολουθούμε τα παρακάτω βήματα:

   Βήμα 1. Ορίζουμε πολυώνυμα βαθμού ανάλογα με την επιλεγόμενη μορφή στοιχείων.

   Βήμα 2. Ορίζουμε ότι τα πολυώνυμα έχουν την ιδιότητα ότι σε ένα σημείο του στοιχείου λαμβάνουν ως τιμή την μονάδα και στα υπόλοιπα σημεία του στοιχείου είναι μηδενικά.

   Βήμα 3. Για να υπολογίσουμε τους πίνακες συντελεστών των πολυωνύμων κάθε στοιχείου οι οποίοι συναντώνται στην βιβλιογραφία ως τοπικός πίνακας δυσκαμψίας (local stifness matrix) και τοπικό διάνυσμα φορτίου (local load vector) και περιέχουν την συνεισφορά κάθε στοιχείου, χρησιμοποιούμε:

   - Την ασθενή μορφή και τις συναρτήσεις B – splines 1ου βαθμού για την μέθοδο Galerkin.
   - Το κριτήριο της πρώτης παραγώγου και τις συναρτήσεις hat για την μέθοδο Ritz.

   Βήμα 4. Ακολουθούμε την μέθοδο δόμηση στοιχείου προς στοιχείο (element by element) για να κατασκευάσουμε τον καθολικό πίνακα δυσκαμψίας (global stifness matrix) και το καθολικό διάνυσμα φορτίου (global load vector).

Ενότητα Δ – Υπολογισμός λύσης

Εφόσον έχουμε καθολικό πίνακα δυσκαμψίας A και το καθολικό διάνυσμα φορτίου F, λύνουμε το σύστημα $AU = F$ για να υπολογίσουμε την προσέγγιση της λύσης.

2.8.1. Τύποι πεπερασμένων στοιχείων

Παρακάτω παραθέτουμε τους βασικούς τύπους των στοιχείων που χρησιμοποιούνται ο δε τρόπος χρήσης του επεξηγείται αναλυτικά στα αντίστοιχα τμήματα των μεθόδων και της αντίστοιχης υλοποίησης τους στη Matlab.

Γραμμικά στοιχεία (linear elements): Τα πολυώνυμα ορίζονται σε κάθε τρίγωνο στο πλέγμα τα οποία απαρτίζουν το σύνολο των συναρτήσεων βάσης του τριγωνικού στοιχείου. Κάθε συνάρτηση βάσης λαμβάνει την τιμή ένα σε έναν κόμβο και μηδεν στους άλλους δηλαδή :

$$\varphi_i(x_i, y_i) = \delta_i^j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

Όταν χρησιμοποιούμε γραμμικά στοιχεία οι συναρτήσεις βάσης είναι γραμμικά πολυώνυμα πρώτου βαθμού, δηλαδή :

$$\varphi_i(x, y) = \alpha_i + b_i x + c_i y$$

Κάθε συνάρτηση $\varphi_i(x, y)$ έχει τρεις άγνωστους συντελεστές τους α_i, b_i και c_i συνεπώς για τις τρεις κορυφές κάθε πεπερασμένου στοιχείου για κάθε συνάρτηση $\varphi_i(x, y)$ παράγονται τρεις εξισώσεις:

$$\begin{array}{lll}
\varphi_1(x_1, y_1) = \alpha_1 + b_1 x_1 + c_1 y_1 = 1 & \varphi_2(x_1, y_1) = \alpha_2 + b_2 x_1 + c_2 y_1 = 0 & \varphi_3(x_1, y_1) = \alpha_3 + b_3 x_1 + c_3 y_1 = 0 \\
\varphi_1(x_2, y_2) = \alpha_1 + b_1 x_2 + c_1 y_2 = 0 & \varphi_2(x_2, y_2) = \alpha_2 + b_2 x_2 + c_2 y_2 = 1 & \varphi_3(x_2, y_2) = \alpha_3 + b_3 x_2 + c_3 y_2 = 0 \\
\varphi_1(x_3, y_3) = \alpha_1 + b_1 x_3 + c_1 y_3 = 0 & \varphi_2(x_3, y_3) = \alpha_2 + b_2 x_3 + c_2 y_3 = 0 & \varphi_3(x_3, y_3) = \alpha_3 + b_3 x_3 + c_3 y_3 = 1
\end{array}$$

Τα παραπάνω συστήματα με μορφή πινάκων ανάγονται στις εξής εξισώσεις :

$$\begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix}\begin{bmatrix} \alpha_1 \\ b_1 \\ c_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix}\begin{bmatrix} \alpha_2 \\ b_2 \\ c_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix}\begin{bmatrix} \alpha_3 \\ b_3 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Συνεπώς για τις τρεις συναρτήσεις βάσεις για το κάθε στοιχείο θα έχουμε ένα σύστημα το οποίο υπό μορφή πινάκων μας δίνεται ως εξής :

$$\begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix}\begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Στο σημείο αυτό να σημειώσουμε ότι ο πίνακας:

$$\begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$$

, που προκύπτει από την λύση του παραπάνω συστήματος για κάθε στοιχείο θα είναι τμήμα του καθολικού πίνακα A που προαναφέραμε και χρησιμοποιείται για την λύση του $AU = F$.

Τετραγωνικά στοιχεία (Quadratic elements): Τα πολυώνυμα ορίζονται στα τρίγωνα που έχουμε ορίσει στο πλέγμα και τα οποία απαρτίζουν το σύνολο των συναρτήσεων βάσης του τριγωνικού στοιχείου, με κάθε συνάρτηση βάσης λαμβάνει την τιμή ένα σε έναν κόμβο και μηδεν στους άλλους δηλαδή:

$$\varphi_i(x_i, y_i) = \delta_i^j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

Όταν χρησιμοποιούμε τετραγωνικά στοιχεία οι συναρτήσεις βάσης είναι γραμμικά πολυώνυμα δευτέρου βαθμού, δηλαδή:

$$\varphi_i(x, y) = \alpha_i + b_i x + c_i y + d_i x^2 + e_i xy + f_i y^2$$

Κάθε συνάρτηση $\varphi_i(x, y)$ έχει εξι άγνωστους συντελεστές τους $\alpha_i, b_i, c_i, d_i, e_i$ και f_i συνεπώς στις εξι κορυφές του πεπερασμένου στοιχείου για κάθε συνάρτηση $\varphi_i(x, y)$ παράγονται εξι εξισώσεις (αναγράφουμε μόνο την περίπτωση της $\varphi_1(x, y)$) :

$$\begin{aligned}
\varphi_1(x_1, y_1) &= \alpha_1 + b_1 x_1 + c_1 y_1 + d_1 x_1^2 + e_1 x_1 y_1 + f_1 y_1^2 = 1 \\
\varphi_1(x_2, y_2) &= \alpha_1 + b_1 x_2 + c_1 y_2 + d_1 x_2^2 + e_1 x_2 y_2 + f_1 y_2^2 = 0 \\
\varphi_1(x_3, y_3) &= \alpha_1 + b_1 x_3 + c_1 y_3 + d_1 x_3^2 + e_1 x_3 y_3 + f_1 y_3^2 = 0 \\
\varphi_1(x_4, y_4) &= \alpha_1 + b_1 x_4 + c_1 y_4 + d_1 x_4^2 + e_1 x_4 y_4 + f_1 y_4^2 = 0 \\
\varphi_1(x_5, y_5) &= \alpha_1 + b_1 x_5 + c_1 y_5 + d_1 x_5^2 + e_1 x_5 y_5 + f_1 y_5^2 = 0 \\
\varphi_1(x_5, y_6) &= \alpha_1 + b_1 x_6 + c_1 y_6 + d_1 x_6^2 + e_1 x_6 y_6 + f_1 y_6^2 = 0
\end{aligned}$$

Το παραπάνω συστήμα με μορφή πινάκων ανάγεται στην εξής εξισώση πινάκων (ανάλογα ισχύουν και για τις υπόλοιπες $\varphi_1(x, y)$):

$$\begin{bmatrix} 1 & x_1 & y_1 & x_1^2 & x_1 y_1 & y_1^2 \\ 1 & x_2 & y_2 & x_2^2 & x_2 y_2 & y_2^2 \\ 1 & x_3 & y_3 & x_3^2 & x_3 y_3 & y_3^2 \\ 1 & x_4 & y_4 & x_4^2 & x_4 y_4 & y_4^2 \\ 1 & x_5 & y_5 & x_5^2 & x_5 y_5 & y_5^2 \\ 1 & x_6 & y_6 & x_6^2 & x_6 y_6 & y_6^2 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ b_1 \\ c_1 \\ d_1 \\ e_1 \\ f_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Συνεπώς για τις εξι συναρτήσεις βάσεις θα έχουμε ένα σύστημα το οποίο υπό μορφή πινάκων μας δίνεται ως εξής:

$$\begin{bmatrix} 1 & x_1 & y_1 & x_1^2 & x_1y_1 & y_1^2 \\ 1 & x_2 & y_2 & x_2^2 & x_2y_2 & y_2^2 \\ 1 & x_3 & y_3 & x_3^2 & x_3y_3 & y_3^2 \\ 1 & x_4 & y_4 & x_4^2 & x_4y_4 & y_4^2 \\ 1 & x_5 & y_5 & x_5^2 & x_5y_5 & y_5^2 \\ 1 & x_6 & y_6 & x_6^2 & x_6y_6 & y_6^2 \end{bmatrix} \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_5 & \alpha_6 \\ b_1 & b_2 & b_3 & b_4 & b_5 & b_6 \\ c_1 & c_2 & c_3 & c_4 & c_5 & c_6 \\ d_1 & d_2 & d_3 & d_4 & d_5 & d_6 \\ e_1 & e_2 & e_3 & e_4 & e_5 & e_6 \\ f_1 & f_2 & f_3 & f_4 & f_5 & f_6 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Κυβικά στοιχεία (Cubic elements): Τα πολυώνυμα ορίζονται στα τρίγωνα που έχουμε ορίσει στο πλέγμα και τα οποία απαρτίζουν το σύνολο των συναρτήσεων βάσης του τριγωνικού στοιχείου, με κάθε συνάρτηση βάσης να λαμβάνει την τιμή ένα σε έναν κόμβο και μηδεν στους άλλους δηλαδή:

$$\varphi_i(x_i, y_i) = \delta_i^j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

Όταν χρησιμοποιούμε κυβικά στοιχεία οι συναρτήσεις βάσης είναι γραμμικά πολυώνυμα τρίτου βαθμού, δηλαδή:

$$\varphi_i(x, y) = \alpha_i + b_i x + c_i y + d_i x^2 + e_i xy + f_i y^2 + g_i x^3 + h_i x^2 y + i_i xy^2 + j_i y^3$$

Κάθε συνάρτηση $\varphi_i(x, y)$ έχει δέκα άγνωστους συντελεστές τους $\alpha_i, b_i, c_i, d_i, e_i, f_i, g_i, h_i.i_i$ και j_i συνεπώς για τις δέκα κορυφές κάθε πεπερασμένου στοιχείου για κάθε συνάρτηση $\varphi_i(x, y)$ παράγονται δέκα εξισώσεις (για οικονομία χώρου δεν αναγράφουμε τις αντίστοιχες περιπτώσεις, διότι είναι κατ' αναλογία με τις προηγούμενες περιπτώσεις στοιχείων).

2.9. Σύνδεση θεωρίας και μεθόδων με ελλειπτικές εξισώσεις

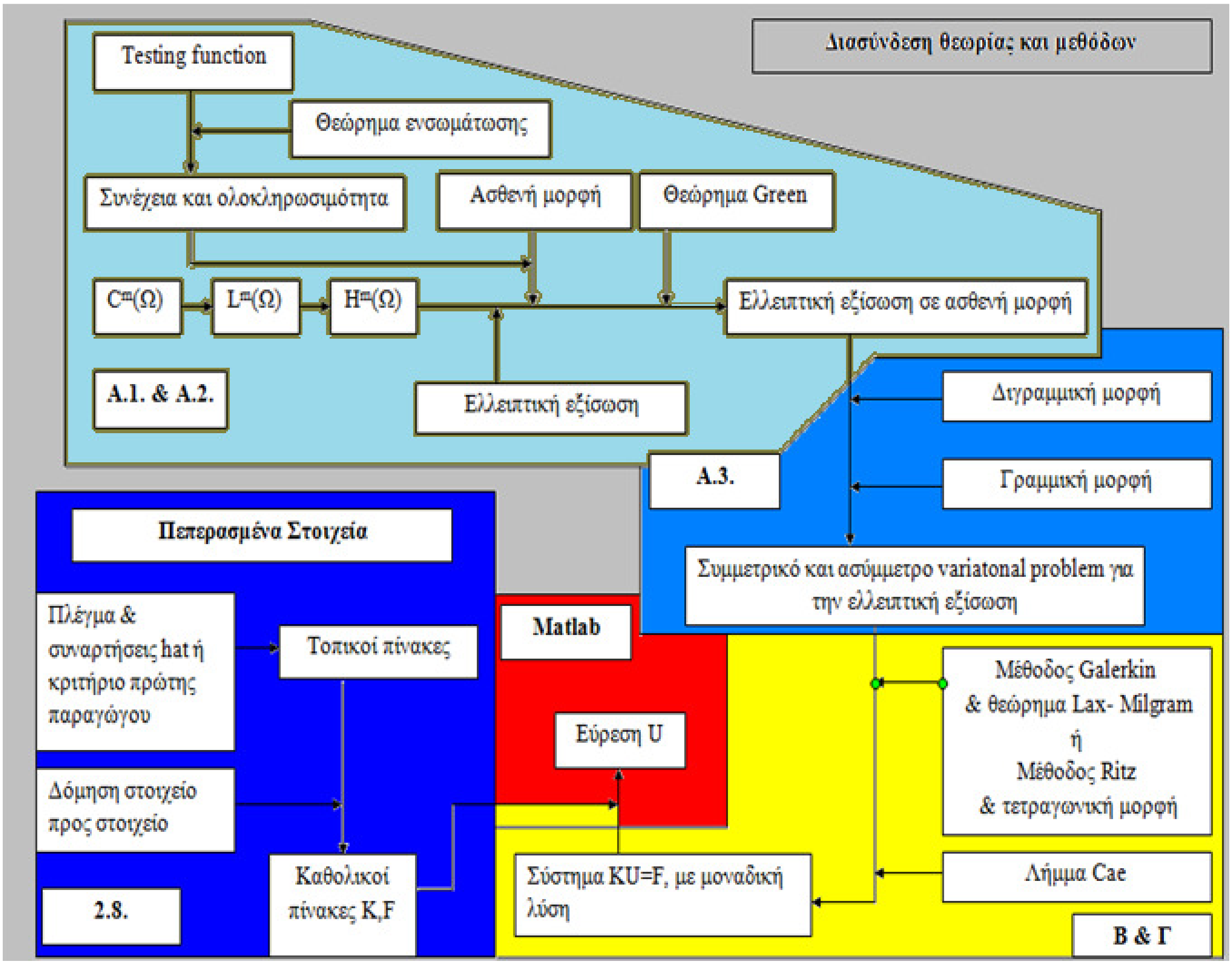
Παρακάτω παραθέτουμε τέσσερεις λογικές ενότητες, οι οποίες συνδέουν την θεωρία και τις μεθόδους που προαναφέραμε, πιο συγκεκριμένα :

Ενότητα Α (Α.1. & Α.2. & Α.3.). Η χρηστικότητα των χώρων Sobolev σχετικά με την συνέχεια και την ολοκληρωσιμότητα συναρτήσεων και η σύνδεση της ελλειπτικής εξίσωσης με την διγραμμική μορφή και κατ' επέκταση με το συμμετρικό και το ασύμμετρο πρόβλημα μεταβολής.

Ενότητα Β. Η σύνδεση της διγραμμικής μορφής με τις μεθόδους και τους πεπερασμένους χώρους.

Ενότητα Γ. Υπάρξη, μοναδικότητα και εγγύηση για το ανω φράγμα της λύσης.

Ενότητα Δ. Σύνοψη και παρατηρήσεις για τις μεθόδους.



Α. Χώροι Sobolev

Η ανάγκη για την χρήση των χώρων Sobolev στην επίλυση εξισώσεων με μερικές παραγώγους πηγάζει από το θεώρημα διαφορισιμότητας [17], όπου γνωρίζουμε ότι εάν οι μερικές παράγωγοι υπάρχουν και είναι συνεχείς σε μια γειτονία του α τότε και η συνάρτηση θα είναι παραγωγίσιμη σε μια γειτονία του α αλλά δεν ισχύει κατ' ανάγκη το αντίστροφο.

Συνεπώς δύναται μια παραγωγίσιμη σε μια γειτονιά συνάρτηση να έχει ασυνεχείς μερικές παραγώγους, π.χ. Η συνάρτηση:

$$f(x,y)=\begin{cases}(x^2+y^2)\sin\left(\dfrac{1}{\sqrt{x^2+y}}\right), & (x,y)\neq(0,0)\\ 0, \quad (x,y)=(0,0)\end{cases}$$

, είναι παραγωγίσιμη αλλά έχει ασυνεχείς και μη ολοκληρώσιμες μερικές παραγώγους [18].

Η χρήση των χώρων Sobolev μας παρέχει δυο βασικά εργαλεία :

- Την αντικατάσταση των συνεχών χώρων με πεπερασμένων διαστάσεων χώρους, όπου μπορούμε να υπολογίσουμε τις συναρτήσεις βάσης και κατ' επέκταση οποιοδήποτε σημείο του χώρου.
- Την δυνατότητα κατασκευής της ασθενούς μορφής χρησιμοποιώντας δοκιμαστικές συναρτήσεις όπου δεν εμφανίζονται τα προβλήμα συνέχειας και ολοκληρωσιμότητας.

A.1. Συνέχεια και ολοκληρωσιμότητα για την ασθενή μορφή

Με την ασθενή μορφή έχουμε ότι η συνάρτηση δοκιμής θα είναι συνεχής και ολοκληρώσιμη, π.χ. έστω η συνάρτηση δοκιμής $v(x,y)\in H^2(\Omega)$, τότε από το θεώρημα ενσωμάτωσης έχουμε:

- $H^2(\Omega)\subset C^0(\Omega)$, συνεπώς $v(x,y)\in C^0(\Omega)$ δηλαδή η συνάρτηση δοκιμής $v(x,y)$ είναι συνεχής.
- $v(x,y)\in H^2(\Omega)\Rightarrow v, v_x, v_y, v_{xx}, v_{yy}\in L^2(\Omega)$, συνεπώς οι μερικές παράγωγοι της δοκιμαστικής συνάρτησης $v(x,y)$ θα είναι ολοκληρώσιμες.

A.2. Σύνδεση της ελλειπτικής εξίσωσης με την ασθενή μορφή

Για να εκμεταλλευτούμε την συνέχεια και την ολοκληρωσιμότητα της συνάρτηση δοκιμής και για να λύσουμε ελλειπτικές διαφορικές εξισώσεις στους χώρους Sobolev θα πρέπει να συνδέσουμε την ασθενή μορφή με την διαφορική εξίσωση, όπου πολλαπλασιάζουμε με μια συνάρτηση δοκιμής, ολοκληρώνουμε επί του Ω, χρησιμοποιούμε την ολοκλήρωση κατά μέρη από το θεώρημα του Green και τέλος αξιοποιούμε το γεγονός ότι η δοκιμαστική συνάρτηση είναι μηδενική επί του συνόρου, π.χ. για την εξίσωση Poisson με μηδενικές συνοριακές συνθήκες στις τρείς διαστάσεις :

$$-\nabla^2 u = f \Leftrightarrow -\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}\right)\cdot v(x,y)=f(x,y)\cdot v(x,y)\Leftrightarrow -\left(u_{xx}(x,y)+u_{yy}(x,y)\right)\cdot v(x,y)=f(x,y)\cdot v(x,y)$$

, θα έχουμε ότι :

$$-\iint_{\Omega}\left(u_{xx}(x,y)+u_{yy}(x,y)\right)\cdot v(x,y)dxdy=\iint_{\Omega}\left(f(x,y)\right)v(x,y)dxdy$$

, δηλαδή :

$$\iint_{\Omega}\left(u_{xx}(x,y)+u_{yy}(x,y)\right)\cdot v(x,y)dxdy=\iint_{\Omega}\left[\left(u_x v\right)_x+\left(u_y v\right)_y-u_x v_x-u_y v_y\right]dxdy \Leftrightarrow$$

$$\iint_{\Omega}\left(u_{xx}(x,y)+u_{yy}(x,y)\right)\cdot v(x,y)dxdy=\iint_{\Omega}\left[\left(u_x v\right)_x+\left(u_y v\right)_y\right]dxdy-\iint_{\Omega}\left[u_x v_x+u_y v_y\right]dxdy \quad (1)$$

Από το θεώρημα του Green γνωρίζουμε ότι :

$$\iint_{\Omega}\left(P_x+Q_y\right)dxdy=\int_{\partial\Omega}(Pdy-Qdx)$$

, συνεπώς έχουμε :

$$\iint_{\Omega}\left[\left(u_x v\right)_x+\left(u_y v\right)_y\right]dxdy=\int_{\partial\Omega}\left(u_x v\right)dy-\left(u_y v\right)dx$$

, όμως η δοκιμαστική είναι μηδενική επί του συνόρου θα είναι μηδέν ή ισοδύναμα ο παραπάνω όρος είναι μηδέν οπότε από την (1), προκύπτει :

$$\iint_{\Omega}\left(u_{xx}(x,y)+u_{yy}(x,y)\right)\cdot v(x,y)dxdy=-\iint_{\Omega}\left[u_x v_x+u_y v_y\right]dxdy \Leftrightarrow$$

$$-\iint_{\Omega}\left(u_{xx}(x,y)+u_{yy}(x,y)\right)\cdot v(x,y)dxdy=\iint_{\Omega}\left[u_x v_x+u_y v_y\right]dxdy \Leftrightarrow$$

$$\iint_{\Omega}\left(f(x,y)\right)v(x,y)dxdy=\iint_{\Omega}\left[u_x v_x+u_y v_y\right]dxdy$$

Εάν χρησιμοποιήσουμε την αποκλίνουσα μορφή, η παραπάνω σχέση γράφεται ως εξής:

$$\iint_{\Omega}\left(fv\right)dx=\iint_{\Omega}\left(\nabla u\cdot\nabla v\right)dx$$

Εν γένει μια γραμμική δεύτερης τάξης διαφορική εξίσωση με μερικές παραγώγους και συντελέστες πραγματικούς αριθμούς, είναι ελλειπτική εάν είναι της μορφής [23] :

$$\alpha\frac{\partial^2 u}{\partial x^2}+b\frac{\partial^2 u}{\partial x\partial y}+c\frac{\partial^2 u}{\partial y^2}+d\frac{\partial u}{\partial x}+e\frac{\partial u}{\partial y}+fu=g,\quad u\in\Omega$$

, με $b^2-4\alpha c<0$ και στην βιβλιογραφία συναντάται ως $Lu=f$, όπου L είναι ο τελεστής :

$$L=-\sum_{i,j=1}^{d}c_{ij}\frac{\partial}{\partial x_i}\left(\frac{\partial}{\partial x_j}\right)+\sum_{i=1}^{d}b_i\frac{\partial u}{\partial x_i}+\alpha u$$

, π.χ. οι εξισώσεις Laplace και Poisson είναι ίσως τα δυο πιο γνωστά παραδείγματα. Κατ' επέκταση το πρόβλημα στην γενική του μορφή είναι το εξής [23] :

$$-\sum_{i,j=1}^{d}c_{ij}\frac{\partial}{\partial x_i}\left(\frac{\partial}{\partial x_j}\right)+\sum_{i=1}^{d}b_i\frac{\partial u}{\partial x_i}+\alpha u=f$$

, το οποίο σε αποκλίνουσα μορφή αναδιατυπώνεται ως εξής [23] :

$$-\left(c(x)\cdot\nabla u\cdot\nabla v\right)+\left(\alpha(x)\cdot u\cdot v\right)=f$$

, όπου :

$$\begin{aligned} &c_{ij}(x)=c(x)\delta_{ij} \\ &b(x)=0, \quad \forall x\in\Omega \end{aligned}$$

Από τα παραπάνω προκύπτει ότι εάν πολλαπλασιάσουμε με μια συνάρτηση δοκιμής και ολοκληρώσουμε επί του πεδίου ορισμού, έχουμε [23] :

$$\int_{\Omega}\sum_{i,j=1}^{d}c_{ij}u_{x_i}u_{x_j}+\alpha uv=\int_{\Omega}fvdx$$

, συνεπώς το πρόβλημα ισοδυναμεί με το να βρεθεί $v\in H^1(\Omega)$ τέτοια ώστε δοθείσας $f\in L^2(\Omega)$ να ισχύει ότι [23] :

$$-\int_{\Omega}\left(c\cdot\nabla u\cdot\nabla v\right)dx+\int_{\Omega}\left(\alpha\cdot u\cdot v\right)dx=\int_{\Omega}\left(fv\right)dx$$

Α.3. Σύνδεση της ασθενούς μορφής με την διγραμμική μορφή και τα προβλήματα μεταβολής

Η χρηστικότητα της ασθενούς μορφής για την ελλειπτική εξίσωσης στους χώρου Sobolev πέραν της συνέχειας και της ολοκληρωσιμότητας, επεκτείνεται στο ότι μπορούμε να την αντιστοιχήσουμε με μια διγραμμική μορφή ως εξής :

- Θεωρούμε συνάρτηση $\alpha(\cdot,\cdot):H^1(\Omega)\text{ x }H^1(\Omega)\to R$ με κανόνα αντιστοίχησης :

$$\alpha(u,v)=\int_{\Omega}\left(c\cdot\nabla u\cdot\nabla v\right)dx+\int_{\Omega}\left(\alpha\cdot u\cdot v\right)dx$$

, η οποία είναι διγραμμική μορφή [23].

- Θεωρούμε συνάρτηση $l\in V'$ με κανόνα αντιστοίχισης :

$$l(u)=<l,v>=\int_{\Omega}\left(f\cdot v\right)dx$$

, με $<\cdot,\cdot>$ να είναι το εσωτερικό γινόμενο στον $L^2(\Omega)$, η οποία είναι γραμμική μορφή [23].

Σύνοψη Α.1. – Α.3.

Από την ελλειπτική εξίσωση μερικών παραγώγων :

$$\alpha\frac{\partial^2u}{\partial x^2}+b\frac{\partial^2u}{\partial x\partial y}+c\frac{\partial^2u}{\partial y^2}+d\frac{\partial u}{\partial x}+e\frac{\partial u}{\partial x}+fu=g, \quad u\in\Omega \ , \ L=-\sum_{i,j=1}^{d}c_{ij}\frac{\partial}{\partial x_i}\left(\frac{\partial}{\partial x_j}\right)+\sum_{i=1}^{d}b_i\frac{\partial}{\partial x_i}+\alpha u$$

, προκύπτει [23] :

$$-\int_{\Omega}\left(c\cdot\nabla u\cdot\nabla v\right)dx+\int_{\Omega}\left(\alpha\cdot u\cdot v\right)dx=\int_{\Omega}\left(fv\right)dx$$

, από όπου εάν θέσουμε [23] :

$$\alpha(u,v)=\int_{\Omega}\left(c\cdot\nabla u\cdot\nabla v\right)dx+\int_{\Omega}\left(\alpha\cdot u\cdot v\right)dx \quad \wedge \quad l(u)=<l,v>=\int_{\Omega}\left(f\cdot v\right)dx$$

, προκύπτει ότι $\alpha(u,v)=l(v),\quad \forall v\in V$ όπου για την συνέχεια, την ολοκληρωσιμότητα, την διγραμμική μορφή και την γραμμική μορφή ισχύει ότι προαναφέραμε στα A1 έως και A3. Συνεπώς το **πρόβλημα επίλυσης της διαφορικής εξίσωση υπό το πρίσμα της ασθενούς μορφής** ανάγεται στο πρόβλημα εύρεσης μιας συνάρτησης , τέτοιας ώστε :

$$\alpha(u,v)=l(v),\quad \forall v\in V$$

, **από όπου πηγάζουν οι ορισμοί του συμμετρικού και ασύμμετρου** πρόβληματος που προαναφέραμε.

B. Σύνδεση της διγραμμικής μορφής με τις μεθόδους

Από τα προαναφερθέντα προκύπτει ότι το πρόβλημα επίλυσης της αρχικής εξίσωση με την επιλογή κατάλληλων διγραμικών μορφών ισοδυναμεί με δυο προβλήματα για τα οποία γνωρίζουμε ότι έχουμε μοναδική λύση, συνεπώς χρησιμοποιούμε τις μεθόδους Ritz και Galerkin έτσι ώστε να έχουμε μετάβαση σε πεπερασμένους χώρους στους οποίους μπορούμε να χρησιμοποιήσουμε εργαλεία της αριθμητικής ανάλυσης. Όπως προαναφέραμε και στις δυο μεθόδους αναζητούμε λύση στον πεπερασμένο χώρο $V_H=\left\{u_h(x),\ \ u_h(x)=\sum_{j=1}^{N}c_j\varphi_j\right\}$ και εφόσον ο τελεστής L είναι γραμμικός ισχύει ότι:

$$Lu_h=L\left(\sum_{j=1}^{N}c_j\varphi_j\right)=\sum_{j=1}^{N}c_jL\varphi_j$$

Για να υπολογίσουμε τους N στο πλήθος αγνώστους c_j χρειαζόμαστε n στο πλήθος εξισώσεις και προκύπτει ότι το σύνολο $\text{span}\{\varphi_1,\dots,\varphi_n\}$ παράγει το ζητούμενο σύστημα εξισώσεων, διότι εφόσον οι φ_i είναι γραμμικά ανεξάρτητες θα έχουμε ότι :

$$\alpha_1\varphi_1+\dots+\alpha_n\varphi_n=0\Rightarrow\alpha_1=\dots=\alpha_n=0$$

, συνεπώς εάν έχουμε :

$$\begin{gathered}\varphi_1\left(c_1Lu_1+c_2Lu_2+\dots+c_nLu_n-f\right)=0\\ \vdots\\ \varphi_n\left(c_1Lu_1+c_2Lu_2+\dots+c_nLu_n-f\right)=0\end{gathered}$$

, τότε προκύπτει :

$$\varphi_1\left(c_1Lu_1+c_2Lu_2+\dots+c_nLu_n-f\right)+\dots+\varphi_n\left(c_1Lu_1+c_2Lu_2+\dots+c_nLu_n-f\right)=0$$

, ή ισοδύναμα:

$$\left.\begin{array}{l} c_1Lu_1 + c_2Lu_2 + ... + c_nLu_n - f = 0 \\ \vdots \\ c_1Lu_1 + c_2Lu_2 + ... + c_nLu_n - f = 0 \end{array}\right\} \Rightarrow \begin{array}{l} c_1Lu_1 + c_2Lu_2 + ... + c_nLu_n = f \\ \vdots \\ c_1Lu_1 + c_2Lu_2 + ... + c_nLu_n = f \end{array}$$

Εφόσον $\sum_{j=1}^{N} c_j Lu_j = L\left(\sum_{j=1}^{N} c_j u_j\right) = Lu$, η παραπάνω σχέση αναδιατυπώνεται ως εξής:

$$Lu = f$$

, που είναι το αρχικό σύστημα που θέλαμε να επιλύσουμε. Από τα παραπάνω προκύπτει, ότι εάν βρούμε ένα σύνολο γραμμικά ανεξάρτητων $\{\varphi_1, ..., \varphi_n\}$, τότε μπορούμε να αντιμετωπίσουμε το αρχικό πρόβλημα $Lu = f$ και για την εύρεση του, παρατηρούμε ότι:

$$\left.\begin{array}{l} \varphi_1\left(c_1Lu_1 + c_2Lu_2 + ... + c_nLu_n - f\right) = 0 \\ \vdots \\ \varphi_n\left(c_1Lu_1 + c_2Lu_2 + ... + c_nLu_n - f\right) = 0 \end{array}\right\} \Rightarrow \left.\begin{array}{l} \varphi_1\left(c_1Lu_1\right) + \varphi_1\left(c_2Lu_2\right) + ... + \varphi_1\left(c_nLu_n\right) = \varphi_1 f \\ \vdots \\ \varphi_n\left(c_1Lu_1\right) + \varphi_n\left(c_2Lu_2\right) + ... + \varphi_n\left(c_nLu_n\right) = \varphi_n f \end{array}\right\} \Rightarrow AC = B$$

με:

$$A = \begin{bmatrix} \varphi_1 Lu_1 & \varphi_1 Lu_2 & \cdots & \varphi_1 Lu_n \\ \varphi_2 Lu_1 & \varphi_2 Lu_2 & \cdots & \varphi_2 Lu_n \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_n Lu_1 & \varphi_n Lu_2 & \cdots & \varphi_n Lu_n \end{bmatrix}, \quad b = \begin{bmatrix} \varphi_1 f \\ \varphi_2 f \\ \vdots \\ \varphi_n f \end{bmatrix}, \quad c = [c_1, ..., c_n]^T$$

, συνεπώς το πρόβλημα ανάγεται στην λύση του συστήματος $AC = B$.

Γ. Υπάρξη, μοναδικότητα και ποιότητα της προσέγγισης

Στα αντίστοιχα τμήματα των μεθόδων αναλύσαμε σχετικά για την ύπαρξη και την μοναδικότητα της λύσης του συστήματος, θεωρήσαμε όμως σκόπιμο αφενός να αναφέρουμε την γενική περίπτωση αφετέρου να εξετάσουμε την ποιότητα της προσέγγισης.

Γ.1. Ύπαρξη και μοναδικότητα της λύσης

Το σύστημα στο οποίο καταλήξαμε ότι πρέπει να λύσουμε έχει την μορφή:

$$\left.\begin{array}{l} \varphi_1\left(c_1Lu_1\right) + \varphi_1\left(c_2Lu_2\right) + ... + \varphi_1\left(c_nLu_n\right) = \varphi_1 f \\ \vdots \\ \varphi_n\left(c_1Lu_1\right) + \varphi_n\left(c_2Lu_2\right) + ... + \varphi_n\left(c_nLu_n\right) = \varphi_n f \end{array}\right\}$$

ή ισοδύναμα :

$$\varphi_i\left[\sum_{j=1}^{N} c_j L u_j - f\right]=0, \quad i=1,...,n$$

Εάν θέσουμε ως $\Phi_i(u) = < u, u_i >$ να είναι το εσωτερικό γινόμενο της u με μια συνάρτηση δοκιμής u_i κατάλληλη έτσι ώστε να τηρούνται οι προυποθέσεις της πράξης, τότε το ερώτημα σχετικά με την ύπαρξη και την μοναδικότητα της λύσης μετατοπίζεται στην εξίσωση [14]:

$$\sum_{j=1}^{N} c_j \left\langle L u_j, u_i \right\rangle = \left\langle f, u_i \right\rangle, \quad i=1,...,n$$

Εφόσον το εσωτερικό γινόμενο στο R είναι μια θετικά ορισμένη συμμετρική διγραμμική μορφή [14], από το θεώρημα Lax – Milgramm προκύπτει ότι υπάρχει μοναδική λύση στην εξίσωση $\alpha(u,v) = l(v), \quad \forall v \in H$ ή ισοδύναμα θα υπάρχει μοναδική λύση στο αρχικό πρόβλημα.

Γ.2. Ακρίβεια προσέγγισης

Λήμμα Cea [16] : Έστω η διγραμμική μορφή $\alpha : VxV \to R$, τέτοια ώστε:

1. Να είναι φραγμένη, δηλαδή υπάρχει σταθερά $C > 0$ τέτοια ώστε : $\alpha(u,v) \le C\|u\|\|v\|, \ \forall u \in H, \forall v \in H$.

2. Να είναι k – ελλειπτική, δηλαδή υπάρχει σταθερά $k > 0$, τέτοια ώστε : $k\|u\|^2 \le \alpha(u,u)$.

Εάν η $u \in V$ και $u_h \in V_h$ είναι οι μοναδικές λύσεις των εξισώσεων $\alpha(u,v) = l(v), \quad \forall v \in V$ και $\alpha(u_h, v_h) = l(v_h), \quad \forall v_h \in V_h$ αντίστοιχα, τότε ισχύει ότι :

$$\|u - u_h\|_V \le \frac{C}{k} \inf_{u_h \in V_h} \|u - u_h\|_V$$

Από το παραπάνω λήμμα προκύπτει, ότι όταν ισχύουν οι προυποθέσεις του θεωρήματος Lax – Milgram τότε η ακρίβεια της προσέγγισης της $u \in V$ από την $u_h \in V_h$ είναι ισοδύναμα ακριβής με την βέλτιση προσέγγιση της $u \in V$ από την μια συνάρτηση που ανήκει στην V_h .

Δ. Σύνοψη και παρατηρήσεις για τις μεθόδους

Συνοψίζοντας τα παραπάνω η κυρίαρχη συλλογιστική που ακολουθείται στους χώρους Sobolev και η σύνδεση τους με τα πεπερασμένα στοιχεία είναι η εξής :

1. Ορίζουμε μια συνάρτηση δοκιμής στον χώρο Sobolev και παράγουμε την ασθενή μορφή για την οποία γνωρίζουμε ότι είναι συνεχής, έχει ολοκληρώσιμες μερικές παραγώγους και συνδέεται με την διαφορική εξίσωση. Μετέπειτα συσχετίζουμε την ασθενή μορφή με κατάλληλη διγραμμική μορφή συνεπώς το αρχικό πρόβλημα της εξίσωσης μετατοπίστηκε στο πρόβλημα επίλυσης του

$\alpha(u,v) = l(v), \quad \forall v \in V$, δηλαδή ανάγεται στην επίλυση του συμμετρικού και του ασύμμετρου variatonal problem όπως αναλύσαμε στα Α1 έως Α3 και στην σύνοψη τους.

2. Για την λύση των προβλημάτων, χρησιμοποιούμε την ασθενή μορφή και τις μεθόδους Galerkin και Ritz όπου παράγουμε ένα σύστημα εξισώσεων για το οποίο λόγω της σύνδεσης με την διγραμμική μορφή, έχουμε εγγύηση για την μοναδικότητα και την ακρίβεια της προσέγγισης της λύσης λόγω των αντίστοιχων θεωρημάτων και του λήμματος Cae.
3. Υπολογίζουμε την λύση του συστήματος χρησιμοποιώντας πεπερασμένα στοιχεία και η λύση είναι οι συντελεστές για την βάση του χώρου, συνεπώς μπορούμε να εκφράσουμε κάθε σημείο ως γραμμικό συνδυασμό των στοχείων της.
4. Εκφράζουμε ένα σύνολο σημείων βάσει των συναρτήσεων βάσης και εκτιμούμε το σφάλμα και τις σχετικές νόρμες.

Στο σημείο αυτό θεωρούμε σκόπιμο να δώσουμε ένα παράδειγμα εφαρμογής έως και το σημείο υλοποίησης των πεπερασμένων στοιχείων, όπου επισέρχεται το υπολογιστικό κομμάτι και η υλοποίηση στη Matlab που αναλύεται στα αντίστοιχα τμήματα παρακάτω. Για την εξίσωση επί του $\Omega = [0,1]$:

$$\begin{aligned} -u'' + bu' + \alpha u &= f \\ u(0) = u(1) &= 0 \end{aligned}$$

, με $u:[0,1] \to R$ η ασθενής μορφή προκύπτει εάν πολλαπλασιάσουμε με μια συνάρτηση δοκιμής $v \in H^1([0,1])$ και ολοκληρώσουμε επί του πεδίου ορισμού Ω, συνεπώς έχουμε:

$$\int_0^1 u'v' + bu'v + \alpha uv = \int_0^1 fu, \quad \forall v \in H^1([0,1])$$

Θέτουμε :

$$\alpha(u,v) = \int_0^1 u'v' + bu'v + \alpha uv, \quad \forall u,v \in H^1([0,1])$$

$$l(v) = \int_0^1 fu, \quad \forall u \in H^1([0,1])$$

Για τις $\alpha(u,v)$ και $l(v)$ έχουμε ότι είναι διγραμμική και γραμμική μορφή αντίστοιχα [12], όμως για να ικανοποιείται η συνθήκη της k-ελλειπτικότητας απαιτούμε περαιτέρω $\alpha(x) - \frac{1}{2}b'(x) \geq 0$ [12]. Συνεπώς προκύπτει ότι όταν $\alpha(x) - \frac{1}{2}b'(x) \geq 0$, η εξίσωση έχει μοναδική ασθενή λύση στο $\forall u \in H^1([0,1])$.

Στο σημείο αυτό θεωρούμε χρήσιμες τις παρακάτω παρατηρήσεις επί των μεθόδων :

- Για το συμμετρικό πρόβλημα, η απόδειξη της ύπαρξης και της μοναδικότητας της λύσης που αναφέρεται στην σχετική παραπομπή στην βιβλιογραφία χρησιμοποιεί την πρόταση Α και το θεώρημα αναπαράστασης Riesz και μας οδηγεί στην μέθοδο Ritz.

- Για το ασύμμετρο πρόβλημα, η απόδειξη της ύπαρξης και της μοναδικότητας της λύσης που αναφέρεται στην σχετική παραπομπή στην βιβλιογραφία χρησιμοποιεί την πρόταση Α και το θεώρημα Lax - Milgram και μας οδηγεί στην μέθοδο Galerkin.

- Εφόσον όταν εκπληρούνται οι συνθήκες του συμμετρικού προβλήματος εκπληρούνται και οι συνθήκες του θεωρήματος Lax - Milgram θα έχουμε ότι όταν η διγραμμική μορφή είναι συμμετρική οι μέθοδοι ταυτίζονται όπως θα δείξουμε και στο αντίστοιχο παράδειγμα της υλοποίησης της μεθόδου Galerkin.

- Εφόσον όταν εκπληρούνται οι συνθήκες του συμμετρικού προβλήματος εκπληρούνται και οι συνθήκες του θεωρήματος Lax - Milgram θα έχουμε ότι όταν η διγραμμική μορφή είναι συμμετρική, τότε ισχύει το λήμμα Cae.

- Η μέθοδος Galerking προτιμάται από την μέθοδο Ritz διότι:
  - Η μέθοδος Ritz απαιτεί την συμμετρία στην διγραμμική μορφή, συνεπώς η μέθοδος Galekrin αντιμετωπίζει μεγαλύτερο εύρος προβλημάτων και θεωρείται γενίκευση της Ritz [14].
  - Η μέθοδος Ritz έχει ως βάση την ελαχιστοποίηση της συνάρτησης $I(\Phi) = \int_S F(x, y, \Phi, \Phi_x, \Phi_y) ds$, η οποία δεν υπάρχει κατ' ανάγκη , π.χ. Σε μη αυτοσυζυγείς εξισώσεις, δηλαδή όταν εμφανίζονται περιττού βαθμού παράγωγοι [22].
  - Η εφαρμογή της μεθόδου Ritz εμφανίζει δυσκολίες σε πολύπλοκες γεωμετρίες και παρουσιάζει πολύ φτωχά αποτελέσματα σε μη γραμμικά προβλήματα [22].
  - Ο πίνακας Α που προκύπτει για την λύση του συστήματος στην μέθοδο Ritz είναι πυκνός σε αντίθεση με την μέθοδο Galerkin όπου είναι αραιός, συνεπώς υπολογιστικά η μέθοδος Galerkin είναι πολύ πιο συμφέρουσα.

3. Μέθοδος Ritz

3.1. Εφαρμογή στην εξίσωση $-\frac{\partial}{\partial x}\left(p(x)\cdot\frac{\partial y}{\partial x}\right)+q(x)\cdot y(x)=f(x,y)$

Εστω ότι έχουμε το παρακάτω μοντέλο :

$$-\frac{\partial}{\partial x}\left(p(x)\cdot\frac{\partial y}{\partial x}\right)+q(x)\cdot y(x)=f(x,y),\quad 0\le x\le 1$$

, με συνοριακές συνθήκες $y(0)=y(1)=1$. Μια συνάρτηση $y\in C^2[0,1]$ είναι μοναδική λύση της εξίσωσης εάν και μόνο εάν ελαχιστοποιεί την συνάρτηση [4] :

$$I[u]=\int_0^1\left[p(x)\cdot\left(u'(x)\right)^2+q(x)\cdot\left(u(x)\right)^2-2f(x)u(x)\right]dx\quad (1)$$

Στόχος μας είναι να κατασκευάσουμε έναν πεπερασμένων διαστάσεων χώρο V_H, έτσι ώστε να εκφράζουμε κάθε σημείο του ως γραμμικό συνδυασμό των συναρτήσεων βάσης $\varphi_1,\varphi_2,...,\varphi_{M-1}$ και εφόσον η προσέγγιση $\varphi(x)$ της λύσης $u(x)$ είναι γραμμικός συνδυασμός των φ_i προκύπτει ότι είναι της μορφής $\varphi(x)=\sum_{i=1}^{n}c_i\varphi_i(x,y)$, συνεπώς αναζητούμε $c_1,...,c_n$ έτσι ώστε να ελαχιστοποιείται η παράσταση:

$$I[\varphi]=I\left[\sum_{i=1}^{n}c_i\varphi_i(x,y)\right]$$

Αντικαθιστούμε στην (1) :

$$I[\varphi(x)]=I\left[\sum_{i=1}^{n}c_i\varphi_i(x,y)\right]\Rightarrow$$
$$I[\varphi(x)]=\int_0^1\left[p(x)\cdot\left(\sum_{i=1}^{n}c_i\varphi_i'(x,y)\right)^2+q(x)\cdot\left(\sum_{i=1}^{n}c_i\varphi_i(x,y)\right)^2-2f(x)\left(\sum_{i=1}^{n}c_i\varphi_i(x,y)\right)\right]dx$$

Για την εύρεση του ελαχίστου, χρησιμοποιούμε το κριτήριο της πρώτης παραγώγου :

$$\frac{\partial I}{\partial c_j}[\varphi(x)]=0,\quad j=1,...,n$$

ή ισοδύναμα :

$$\frac{\partial I}{\partial c_j}[\varphi(x)]=\int_0^1\left[2p(x)\cdot\left(\sum_{i=1}^{n}c_i\varphi_i'(x,y)\varphi_j'(x,y)\right)+2q(x)\cdot\left(\sum_{i=1}^{n}c_i\varphi_i(x,y)\varphi_j(x,y)\right)-2f(x)\varphi_j(x,y)\right]dx\Leftrightarrow$$
$$\sum_{i=1}^{n}c_i\left[\int_0^1\left\{p(x)\varphi_i'(x,y)\varphi_j'(x,y)+q(x)\varphi_i(x,y)\varphi_j(x,y)\right\}dx\right]-\int_0^1 f(x)\varphi_j(x,y)dx=0,\quad j=1,...,n$$

Από την παραπάνω εξίσωση παράγεται ένα σύστημα εξισώσεων n x n, διότι για κάθε μια εκ των n εις το πλήθος τιμές της j έχουμε άθροισμα n όρων και σε μορφή πινάκων το σύστημα μπορεί να λάβει την μορφή $AC = B$, όπου :

$$\alpha_{ij} = \int_0^1 \left\{ p(x)\varphi_i'(x,y)\varphi_j'(x,y) + q(x)\varphi_i(x,y)\varphi_j(x,y) \right\} dx, \quad \beta_i = \int_0^1 f(x)\varphi_j(x,y)dx, \quad c = [c_1,...,c_n]^T$$

Η εφαρμογή της παραπάνω μεθοδολογίας χωρίζεται σε δυο σκέλη όπου αρχικά χρησιμοποιούμε τις συναρτήσεις βάσης για να εξάγουμε την μορφή των στοιχείων των πινάκων του συστήματος και μετέπειτα στο αριθμητικό σκέλος όπου υπολογίζουμε τα εμπλεκόμενα ολοκληρηματα με την υλοποίηση στη Matlab.

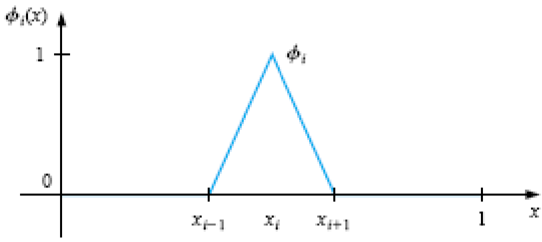
Υπολογισμός των συναρτήσεων βάσης

1ο βήμα : Παράγουμε ενα πλέγμα, $x_0 = 0,...,x_{n+1} = 1$ και ορίζουμε:

- Ως κόμβος ή κομβικό σημείο να είναι τα x_i.
- Ως εκλέπτυνση του πλέγματος (βήμα) να είναι η $h = \max\limits_{0 \le i \le M-1} \{h_i\}$ με $h_i = x_{i+1} - x_i, \quad i = 0,...,n$.

2ο βήμα : Ορίζουμε ότι οι συναρτήσεις βάσης που θα χρησιμοποιήσουμε είναι B – splines 1ου βαθμού, δηλαδή είναι της μορφής [4] :

$$\widehat{B}_i(x) = \varphi_i(x) = \begin{cases} 0 & , \text{εάν } x < x_{i-1} \\ \dfrac{x - x_{i-1}}{h_i} & , \text{εάν } x_{i-1} \le x < x_i \\ \dfrac{x_{i+1} - x}{h_{i+1}} & , \text{εάν } x_i \le x < x_{i+1} \\ 0 & , \text{αλλιώς} \end{cases}$$



Για τις παραγώγους των συναρτήσεων της παραπάνω μορφής για $i = 1,...,n$, θα έχουμε :

$$\varphi_i'(x) = \begin{cases} 0 & , \text{ εάν } x < x_{i-1} \\ \dfrac{1}{h_i} & , \text{ εάν } x_{i-1} \le x < x_i \\ \dfrac{-1}{h_{i+1}} & , \text{ εάν } x_i \le x < x_{i+1} \\ 0 & , \text{ αλλιώς} \end{cases}$$

Από τα παραπάνω προκύπτει ότι οι συναρτήσεις $\varphi_i(x)$ και $\varphi_i'(x)$ θα είναι μη μηδενικές μόνο στο διάστημα (x_{i-1}, x_{i+1}), συνεπώς ισχύει ότι:

$$\varphi_i(x) \cdot \varphi_j(x) = 0 \;\; \wedge \;\; \varphi_i'(x) \cdot \varphi_j'(x) = 0, \quad \forall j \ne i-1, i, i+1$$

3ο βήμα : Βάσει των παραπάνω, για τα στοιχεία των πινάκων Α και του πίνακα Β, προκύπτουν τα εξής :

Πίνακας Α

$$\alpha_{i,i} = \int_0^1 \left\{ p(x)\varphi_i^{'}(x,y)\varphi_i^{'}(x,y) + q(x)\varphi_i(x,y)\varphi_i(x,y) \right\} dx \Leftrightarrow$$

$$\alpha_{i,i} = \int_{x_{i-1}}^{x_i} p(x)\left(\frac{1}{h_i}\right)^2 dx + q(x)\left(\frac{x-x_{i-1}}{h_i}\right)^2 dx + \int_{x_i}^{x_{i+1}} p(x)\left(\frac{-1}{h_{i+1}}\right)^2 dx + \left(\frac{x_{i+1}-x}{h_{i+1}}\right)^2 q(x)dx \Leftrightarrow$$

$$\alpha_{i,i} = \int_{x_{i-1}}^{x_i} p(x)\left(\frac{1}{h_i}\right)^2 dx + \int_{x_i}^{x_{i+1}} p(x)\left(\frac{-1}{h_i}\right)^2 dx + \int_{x_{i-1}}^{x_i} q(x)\left(\frac{1}{h_{i+1}}\right)^2 \left(x-x_{i-1}\right)^2 dx + \int_{x_i}^{x_{i+1}} q(x)\left(\frac{1}{h_{i+1}}\right)^2 \left(x_{i+1}-x\right)^2 dx, \quad i = 1,...n$$

$$\alpha_{i,i+1} = \int_0^1 \left\{ p(x)\varphi_i^{'}(x,y)\varphi_{i+1}^{'}(x,y) + q(x)\varphi_i(x,y)\varphi_{i+1}(x,y) \right\} dx \Leftrightarrow$$

$$\alpha_{i,i+1} = \int_{x_i}^{x_{i+1}} \left(\frac{-1}{h_{i+1}}\right)\left(\frac{1}{h_{i+1}}\right) p(x)dx + \int_{x_i}^{x_{i+1}} \left(\frac{x_{i+1}-x}{h_{i+1}}\right)\left(\frac{x-x_i}{h_{i+1}}\right) q(x)dx \Leftrightarrow$$

$$\alpha_{i,i+1} = \int_{x_i}^{x_{i+1}} -\left(\frac{1}{h_{i+1}}\right)^2 p(x)dx + \int_{x_i}^{x_{i+1}} \left(\frac{1}{h_{i+1}}\right)^2 \left(x_{i+1}-x\right)\left(x-x_i\right) q(x)dx, \quad i = 1,...n-1$$

$$\alpha_{i,i-1} = \int_0^1 \left\{ p(x)\varphi_i^{'}(x,y)\varphi_{i-1}^{'}(x,y) + q(x)\varphi_i(x,y)\varphi_{i-1}(x,y) \right\} dx \Leftrightarrow$$

$$\alpha_{i,i-1} = \int_{x_{i-1}}^{x_i} \left(\frac{1}{h_i}\right)\left(-\frac{1}{h_i}\right) p(x)dx + \int_{x_{i-1}}^{x_i} \left(\frac{x-x_{i-1}}{h_i}\right)\left(\frac{x_i-x}{h_i}\right) q(x)dx, \quad i = 2,...n$$

$$\alpha_{i,i-1} = \int_{x_{i-1}}^{x_i} -\left(\frac{1}{h_{i-1}}\right)^2 p(x)dx + \int_{x_{i-1}}^{x_i} \left(\frac{1}{h_i}\right)^2 \left(x_i-x\right)\left(x-x_{i-1}\right) q(x)dx, \quad i = 2,...n$$

Πίνακας Β

$$\beta_i = \int_0^1 f(x)\varphi_i(x,y)dx \Leftrightarrow$$

$$\beta_i = \int_{x_{i-1}}^{x_i} \left(\frac{x_i-x_{i-1}}{h_i}\right) f(x)dx + \int_{x_i}^{x_{i+1}} \left(\frac{x_{i+1}-x}{h_i}\right) f(x)dx, \quad i = 1,...n$$

Λόγω του απαιτούμενου όγκου στους υπολογισμούς η αριθμητική εκτίμηση των παραπάνω που είναι και το δεύτερο σκέλος έγινε με τον σχετικό κώδικα στη Matlab που παρατίθεται στο παράρτημα τα δε αποτελέσματα και η μεθοδολογία απόκτησης χαμηλών σφαλμάτων επεξηγούνται πλήρως στην παράγραφο 3.3. όπου τροποιήσαμε και επεκτείναμε την αρχική μορφή του σχετικού ψευδοκώδικα [4].

3.2. Εφαρμογή στην εξίσωση $\frac{\partial}{\partial x}\left(p(x,y)\frac{\partial u}{\partial x}\right)+\frac{\partial}{\partial y}\left(q(x,y)\frac{\partial u}{\partial y}\right)+r(x,y)(x,y)=f(x,y)$

Όπως προαναφέραμε στο δεύτερο κεφάλαιο και θεωρούμε εύληπτο από την ανάλυση του παραδείγματος της παραγράφου 3.1. , η εφαρμογή της μέθοδου Ritz έχει μεγάλο κόστος και όπως θα διαφανεί και στο τμήμα των πειραματικών αποτελεσμάτων για να αποκτηθούν υψηλής ποιότητα αποτελέσματα για το σφάλμα, απαιτείται μεγάλο πλήθος επαναλήψεων και πολύ μικρό βήμα.

Παρόλα ταύτα και για λόγους πληρότητας, θεωρήσαμε σκόπιμο να αναλύσουμε ένα πιο πολύπλοκο μοντέλο, έως και το σημείο της παραγωγής της εξίσωσης του συστήματος τα δε περαιτέρω βήματα τόσο στην παραγωγή των ολοκληρωμάτων όσο και στην υλοποίηση στον κώδικα είναι ανάλογα με αυτά που ακολουθήσαμε παραπάνω. Έστω το παρακάτω μοντέλο μερικών διαφορικών εξισώσεων:

$$\frac{\partial}{\partial x}\left(p(x,y)\cdot\frac{\partial u}{\partial x}\right)+\frac{\partial}{\partial y}\left(q(x,y)\cdot\frac{\partial u}{\partial y}\right)+r(x,y)\cdot u(x,y)=f(x,y) \quad (1)$$

, με $(x,y)\in D$ όπου D είναι μια περιοχή στο επίπεδο, με σύνορο S και ότι το σύνολο των συνοριακών συνθηκών απαρτίζεται από τα σύνολα:

- S_1, όπου για την $u(x,y)$ ικανοποιούνται συνοριακές συνθήκες της μορφής: $u(x,y)=g(x,y)$ (2)
- S_2, όπου είναι η υπόλοιπη περιοχή του συνόρου και για την $u(x,y)$ ικανοποιούνται συνοριακές συνθήκες της μορφής:

$$p(x,y)\cdot\frac{\partial u}{\partial x}\cdot\cos(\theta_1)+q(x,y)\cdot\frac{\partial u}{\partial y}\cdot\cos(\theta_2)+g_1(x,y)\cdot u(x,y)=g_2(x,y) \quad (3)$$

Για τις συναρτήσεις του μοντέλλου, θεωρούμε ότι :

- Οι p,q και r είναι συνεχείς συναρτήσεις στο $D\cup S,\ S=S_1\cup S_2$.
- Οι g_1 και g_2 έχουν συνεχείς μερικές πρώτες παραγώγους στο S_2.
- Οι $p(x,y)>0, q(x,y)>0, r(x,y)\le 0$ και $g_1(x,y)>0$.

Η λύση της (1) θα είναι αυτή η οποία ελαχιστοποιεί την παράσταση [4] :

$$I[w]=\iint_D\left\{\frac{1}{2}\cdot\left[p(x,y)\cdot\left(\frac{\partial w}{\partial x}\right)^2+q(x,y)\cdot\left(\frac{\partial w}{\partial y}\right)^2-r(x,y)\cdot w^2\right]+f(x,y)\cdot w\right\}dxdy$$
$$+\int_{P_2}\left\{-g_2(x,y)w+\frac{1}{2}g_1(x,y)w^2\right\}dS$$

Η παραπάνω συνάρτηση προκύπτει παράγοντας την ασθενή μορφή από το μοντέλλο, δηλαδή πολλαπλασιάζοντας κατά μέλη με μια συνάρτηση δοκιμής και ολοκληρώνοντας κατά μέλη και μετέπειτα εφαρμόζουμε το αποτέλεσμα ότι το πρόβλημα ανάγεται στην εύρεση $U=(u_1,...,u_N)\in R^N$, τέτοιο ώστε :

$$J(U)=\inf_{VER^N} J(V)$$

Η μέθοδος των πεπερασμένων στοιχείων κατασκευάζει μια προσέγγιση της μορφής $\varphi(x,y)=\sum_{i=1}^{m}\gamma_i\varphi_i(x,y)$, όπου $\varphi_1,\varphi_2,\ldots,\varphi_m$ είναι γραμμικώς ανεξάρτητα τμηματικά πολυώνυμα πρώτου βαθμού και $\gamma_1,\gamma_2,\ldots,\gamma_m$ σταθερές οι οποίες χρησιμοποιούνται ως εξής:

- Οι $\gamma_1,\gamma_2,\ldots,\gamma_n$ χρησιμοποιούνται για να ελαχιστοποιηθεί η παράσταση $I\left[\sum_{i=1}^{m}\gamma_i\varphi_i(x,y)\right]$.
- Οι $\gamma_{n+1},\gamma_{n+2},\ldots,\gamma_m$ χρησιμοποιούνται για να ικανοποιείται η συνοριακή συνθήκη $u(x,y)=g(x,y)$ (2)

Στην μέθοδο Ritz αναζητούμε w που ελαχιστοποιεί την παράσταση:

$$I[w]=\iint_D\left\{\frac{1}{2}\cdot\left[p(x,y)\cdot\left(\frac{\partial w}{\partial x}\right)^2+q(x,y)\cdot\left(\frac{\partial w}{\partial y}\right)^2-r(x,y)\cdot w^2\right]+f(x,y)\cdot w\right\}dxdy$$
$$+\int_{P_2}\left\{-g_2(x,y)w+\frac{1}{2}g_1(x,y)w^2\right\}dS,$$

, συνεπώς θα έχουμε ότι:

$$I[\varphi]=I\left[\sum_{i=1}^{m}\gamma_i\varphi_i(x,y)\right]\Leftrightarrow$$

$$I[\varphi]=\iint_D\left\{\frac{1}{2}\cdot\left[p(x,y)\cdot\left(\sum_{i=1}^{m}\gamma_i\frac{\partial\varphi_i}{\partial x}(x,y)\right)^2+q(x,y)\cdot\left(\sum_{i=1}^{m}\gamma_i\frac{\partial\varphi_i}{\partial y}(x,y)\right)^2-r(x,y)\cdot\left(\left[\sum_{i=1}^{m}\gamma_i\varphi_i(x,y)\right]\right)^2\right]+f(x,y)\cdot\left[\sum_{i=1}^{m}\gamma_i\varphi_i(x,y)\right]\right\}dxdy$$
$$+\int_{P_2}\left\{-g_2(x,y)\left[\sum_{i=1}^{m}\gamma_i\varphi_i(x,y)\right]+\frac{1}{2}g_1(x,y)\left[\sum_{i=1}^{m}\gamma_i\varphi_i(x,y)\right]^2\right\}dS$$

Για να ελαχιστοποιείτε η συνάρτηση ως προς $\gamma_1,\gamma_2,\ldots,\gamma_n$, θα πρέπει:

$$\frac{\partial I}{\partial\gamma_j}=0,\ \ j=1,\ldots,n$$

Γνωρίζουμε ότι:

$$\frac{1}{2}\cdot\frac{\partial}{\partial\gamma_j}\left(p(x,y)\cdot\left(\sum_{i=1}^{m}\gamma_i\frac{\partial\varphi_i}{\partial x}(x,y)\right)^2\right)=2\cdot\frac{1}{2}\cdot p(x,y)\cdot\sum_{i=1}^{m}\gamma_i\frac{\partial\varphi_i}{\partial x}(x,y)\cdot\frac{\partial\varphi_j}{\partial x}(x,y)=p(x,y)\cdot\sum_{i=1}^{m}\gamma_i\frac{\partial\varphi_i}{\partial x}(x,y)\cdot\frac{\partial\varphi_j}{\partial x}(x,y)$$

$$\frac{\partial}{\partial\gamma_j}\left(f(x,y)\cdot\left[\sum_{i=1}^{m}\gamma_i\varphi_i(x,y)\right]\right)=f(x,y)\cdot\varphi_j(x,y)$$

$$\frac{\partial}{\partial\gamma_j}\left(-g_2(x,y)\left[\sum_{i=1}^{m}\gamma_i\varphi_i(x,y)\right]\right)=-g_2(x,y)\cdot\varphi_j(x,y)$$

$$\frac{\partial}{\partial\gamma_j}\left(\frac{1}{2}g_1(x,y)\left[\sum_{i=1}^{m}\gamma_i\varphi_i(x,y)\right]^2\right)=g_1(x,y)\cdot\varphi_j(x,y)\cdot\sum_{i=1}^{m}\gamma_i\varphi_i(x,y)$$

, διότι για $\gamma_k, k\neq j$ η μερική παράγωγος μηδενίζεται. Συνεπώς για την συνθήκη:

$$\frac{\partial I}{\partial\gamma_j}=0,\ \ j=1,\ldots,n\,,$$

θα έχουμε ότι:

$$\begin{aligned}\frac{\partial I}{\partial\gamma_j}=&\iint_D\{p(x,y)\cdot\sum_{i=1}^{m}\gamma_i\frac{\partial\varphi_i}{\partial x}(x,y)\cdot\frac{\partial\varphi_j}{\partial x}(x,y)+q(x,y)\cdot\sum_{i=1}^{m}\gamma_i\frac{\partial\varphi_i}{\partial y}(x,y)\cdot\frac{\partial\varphi_j}{\partial y}(x,y)-r(x,y)\cdot\sum_{i=1}^{m}\gamma_i\frac{\partial\varphi_i}{\partial x}(x,y)\cdot\varphi_j(x,y)\\&+f(x,y)\varphi_j(x,y)\}dxdy\\&+\int_{P_2}\left\{-g_2(x,y)\varphi_j(x,y)+g_1(x,y)\varphi_j(x,y)\left[\sum_{i=1}^{m}\gamma_i\varphi_i(x,y)\right]\right\}dS,\ \ j=1,...,n\end{aligned}$$

Στο σημείο αυτό ομαδοποιούμε τους παράγοντες στους οποίους εμφανίζεται το άθροισμα και λαμβάνουμε:

$$\begin{aligned}\frac{\partial I}{\partial\gamma_j}=&\sum_{i=1}^{m}\gamma_i\left[\iint_D\left(p(x,y)\cdot\frac{\partial\varphi_i}{\partial x}(x,y)\cdot\frac{\partial\varphi_j}{\partial x}(x,y)+q(x,y)\cdot\frac{\partial\varphi_i}{\partial y}(x,y)\cdot\frac{\partial\varphi_j}{\partial y}(x,y)-r(x,y)\cdot\frac{\partial\varphi_i}{\partial x}(x,y)\cdot\varphi_j(x,y)\right)dxdy\right]\\&+\sum_{i=1}^{m}\gamma_i\left[\int_{P_2}\Big(g_1(x,y)\varphi_i(x,y)\varphi_j(x,y)\Big)dS\right]\\&+\iint_D f(x,y)\varphi_j(x,y)dxdy\\&-\int_{P_2}\Big\{g_2(x,y)\varphi_j(x,y))\Big\}dS,\ \ j=1,...,n\end{aligned}$$

Από την προηγούμενη σχέση και από την $\frac{\partial I}{\partial\gamma_j}=0,\ \ j=1,\ldots,n$, θα έχουμε ότι:

$$\begin{aligned}&\sum_{i=1}^{m}\gamma_i\left[\iint_D\left(p(x,y)\cdot\frac{\partial\varphi_i}{\partial x}(x,y)\cdot\frac{\partial\varphi_j}{\partial x}(x,y)+q(x,y)\cdot\frac{\partial\varphi_i}{\partial y}(x,y)\cdot\frac{\partial\varphi_j}{\partial y}(x,y)-r(x,y)\cdot\frac{\partial\varphi_i}{\partial x}(x,y)\cdot\varphi_j(x,y)\right)dxdy\right]+\\&\sum_{i=1}^{m}\gamma_i\left[\int_{P_2}\Big(g_1(x,y)\varphi_i(x,y)\varphi_j(x,y)\Big)dS\right]+\\&\iint_D f(x,y)\varphi_j(x,y)dxdy-\int_{P_2}\Big\{g_2(x,y)\varphi_j(x,y))\Big\}dS=0,\ \ j=1,...,n\end{aligned}$$

Εάν θέσουμε :

$$a_{ij} = \iint_D \left\{ p(x,y) \cdot \frac{\partial \varphi_i}{\partial x}(x,y) \cdot \frac{\partial \varphi_j}{\partial x}(x,y) + q(x,y) \cdot \frac{\partial \varphi_i}{\partial y}(x,y) \cdot \frac{\partial \varphi_j}{\partial y}(x,y) - r(x,y) \cdot \frac{\partial \varphi_i}{\partial x}(x,y) \cdot \varphi_j(x,y) \right\} dxdy$$
$$+ \int_{P_2} g_1(x,y)\varphi_i(x,y)\varphi_j(x,y)dS, \ i = 1,...,n \ \& \ j = 1,...,n$$

$$\beta_i = -\iint_D f(x,y)\varphi_i(x,y)dxdy + \int_{P_2} g_2(x,y)\varphi_i(x,y))dS - \sum_{k=n+1}^{m} a_{ik}\gamma_k,$$
$$, i = 1,...,n$$
$$b = (\beta_1,...,\beta_n)'$$
$$c = (\gamma_1,...,\gamma_n)'$$

, τότε το πρόβλημα ανάγεται στην λύση της εξίσωσης $Ac = b$.

3.3. Πειραματικά αποτελέσματα

Παρακάτω παραθέτουμε τα αποτελέσματα εφαρμογής της μεθόδου Ritz στην εξίσωση $-\frac{\partial}{\partial x}\left(p(x) \cdot \frac{\partial y}{\partial x} \right) + q(x) \cdot y(x) = f(x,y)$ για τη περίπτωση όπου :

$$p(x) = 1, \quad q(x) = \pi^2, \quad f(x,y) = 2\pi^2 \eta\mu(x)$$

και για την βελτιστοποίηση των αποτελεσμάτων που παρήχθησαν στη Matlab κατασκευάσαμε έναν βρόγχο, ο οποίος σε κάθε επανάληψη λειτουργεί ως εξής:

- Το πλήθος των σημείων N, αυξάνει κατά δέκα με αρχική τιμή το δέκα.
- Επανεκτελείται ο αλγόριθμος εύρεσης της λύσης.
- Τα αποτελέσματα σχετικά με το N, την νόρμα άπειρο για το απόλυτο σφάλμα και το ποσοστιαίο σφάλμα προβάλλονται στην οθόνη.
- Τα αποτελέσματα σχετικά με την λύση, την τιμή της συνάρτησης, την τιμή της προσέγγισης, το σφάλμα, το ποσοστιαίο σφάλμα και την νόρμα άπειρο καταγράφονται στο αρχείο Ritz.txt.
- Εάν η τιμή της νόρμα άπειρο λάβει τιμή μικρότερη ή του 10^{-7} διακόπεται η λετουργία του αλγορίθμου.

Η δομή του αρχείου Ritz.txt που παράγεται είναι η εξής:

- Πλήθος σημείων (N).
- Οι τιμές των σημείων (X).
- Η τιμή της συνάρτησης για το σημείο.
- Η εκτίμηση για το σημείο.
- Το απόλυτο σφάλμα για το σημείο.
- Το ποσοστιαίο σφάλμα για το σημείο.
- Η μέγιστη νόρμα για τα απόλυτο σφάλμα.
- Η μέγιστη νόρμα για το ποσοσταίο σφάλμα.

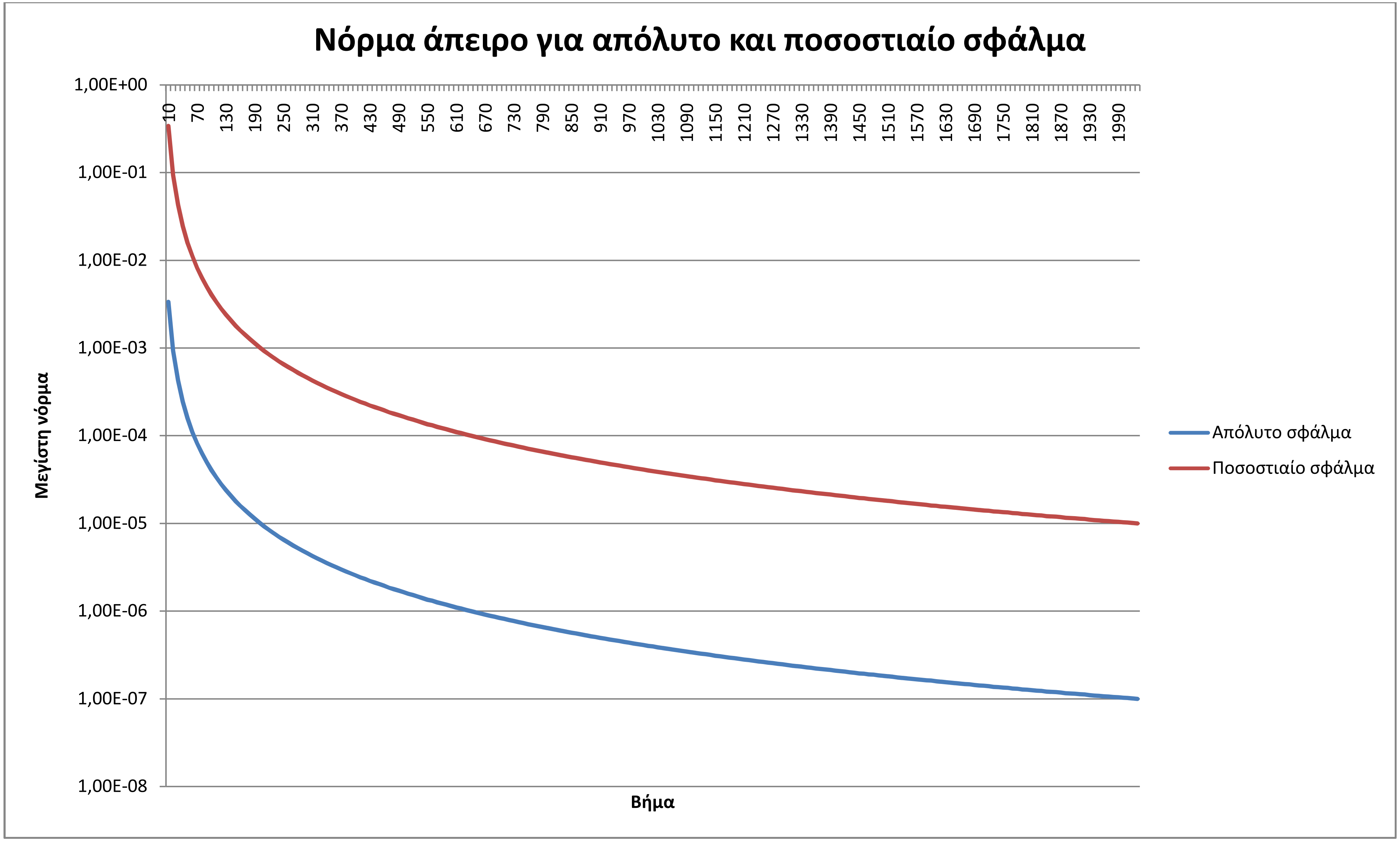
```

---------------------------------------------------------------------------
N =   20
---------------------------------------------------------------------------
X            Πραγματική   Εκτίμηση     Απόλυτο Σφάλμα Ποσοστιαίο Σφάλμα
0.0000000000 0.1490422662 0.1491812173 0.0001389512   0.0932293680120901690
0.0476190476 0.2947551744 0.2950299728 0.0002747984   0.0932293680120800520
0.0952380952 0.4338837391 0.4342882462 0.0004045071   0.0932293680120651890
0.1428571429 0.5633200581 0.5638452378 0.0005251797   0.0932293680120765270
0.1904761905 0.6801727378 0.6808068585 0.0006341207   0.0932293680120624000
0.2380952381 0.7818314825 0.7825603790 0.0007288966   0.0932293680120624140
0.2857142857 0.8660254038 0.8668327938 0.0008073900   0.0932293680120730020
0.3333333333 0.9308737486 0.9317415964 0.0008678477   0.0932293680120758190
0.3809523810 0.9749279122 0.9758368313 0.0009089191   0.0932293680120677290
0.4285714286 0.9972037972 0.9981334840 0.0009296868   0.0932293680120715870
0.4761904762 0.9972037972 0.9981334840 0.0009296868   0.0932293680120827170
0.5238095238 0.9749279122 0.9758368313 0.0009089191   0.0932293680121132760
0.5714285714 0.9308737486 0.9317415964 0.0008678477   0.0932293680121354380
0.6190476190 0.8660254038 0.8668327938 0.0008073900   0.0932293680121242530
0.6666666667 0.7818314825 0.7825603790 0.0007288966   0.0932293680120765690
0.7142857143 0.6801727378 0.6808068585 0.0006341207   0.0932293680120950120
0.7619047619 0.5633200581 0.5638452378 0.0005251797   0.0932293680121356050
0.8095238095 0.4338837391 0.4342882462 0.0004045071   0.0932293680121035470
0.8571428571 0.2947551744 0.2950299728 0.0002747984   0.0932293680119292420
0.9047619048 0.1490422662 0.1491812173 0.0001389512   0.0932293680118851390

Νόρμα άπειρο απολύτων σφαλμάτων =  0.0009296868
Νόρμα άπειρο σχετικων σφαλμάτων =  0.0932293680
---------------------------------------------------------------------------
```

N	Απόλυτο	Ποσοστιαίο	N	Απόλυτο	Ποσοστιαίο	N	Απόλυτο	Ποσοστιαίο	N	Απόλυτο	Ποσοστιαίο	N	Απόλυτο	Ποσοστιαίο
10	3,36E-03	3,40E-01	**430**	2,21E-06	2,21E-04	**850**	5,68E-07	5,68E-05	**1270**	2,55E-07	2,55E-05	**1690**	1,44E-07	1,44E-05
20	9,30E-04	9,32E-02	**440**	2,11E-06	2,11E-04	**860**	5,55E-07	5,55E-05	**1280**	2,51E-07	2,51E-05	**1700**	1,42E-07	1,42E-05
30	4,27E-04	4,28E-02	**450**	2,02E-06	2,02E-04	**870**	5,42E-07	5,42E-05	**1290**	2,47E-07	2,47E-05	**1710**	1,41E-07	1,40E-05
40	2,44E-04	2,45E-02	**460**	1,94E-06	1,94E-04	**880**	5,30E-07	5,30E-05	**1300**	2,43E-07	2,43E-05	**1720**	1,39E-07	1,39E-05
50	1,58E-04	1,58E-02	**470**	1,85E-06	1,85E-04	**890**	5,18E-07	5,18E-05	**1310**	2,39E-07	2,39E-05	**1730**	1,37E-07	1,37E-05
60	1,10E-04	1,11E-02	**480**	1,78E-06	1,78E-04	**900**	5,07E-07	5,07E-05	**1320**	2,36E-07	2,36E-05	**1740**	1,36E-07	1,36E-05
70	8,16E-05	8,16E-03	**490**	1,71E-06	1,71E-04	**910**	4,96E-07	4,96E-05	**1330**	2,32E-07	2,32E-05	**1750**	1,34E-07	1,34E-05
80	6,27E-05	6,27E-03	**500**	1,64E-06	1,64E-04	**920**	4,85E-07	4,85E-05	**1340**	2,29E-07	2,29E-05	**1760**	1,33E-07	1,33E-05
90	4,97E-05	4,97E-03	**510**	1,57E-06	1,57E-04	**930**	4,75E-07	4,74E-05	**1350**	2,25E-07	2,25E-05	**1770**	1,31E-07	1,31E-05
100	4,03E-05	4,03E-03	**520**	1,52E-06	1,52E-04	**940**	4,64E-07	4,64E-05	**1360**	2,22E-07	2,22E-05	**1780**	1,30E-07	1,30E-05
110	3,34E-05	3,34E-03	**530**	1,46E-06	1,46E-04	**950**	4,55E-07	4,55E-05	**1370**	2,19E-07	2,19E-05	**1790**	1,28E-07	1,28E-05
120	2,81E-05	2,81E-03	**540**	1,41E-06	1,41E-04	**960**	4,45E-07	4,45E-05	**1380**	2,16E-07	2,16E-05	**1800**	1,27E-07	1,27E-05
130	2,40E-05	2,40E-03	**550**	1,35E-06	1,35E-04	**970**	4,36E-07	4,36E-05	**1390**	2,13E-07	2,13E-05	**1810**	1,25E-07	1,25E-05
140	2,07E-05	2,07E-03	**560**	1,31E-06	1,31E-04	**980**	4,27E-07	4,27E-05	**1400**	2,10E-07	2,10E-05	**1820**	1,24E-07	1,24E-05
150	1,80E-05	1,80E-03	**570**	1,26E-06	1,26E-04	**990**	4,19E-07	4,19E-05	**1410**	2,07E-07	2,07E-05	**1830**	1,23E-07	1,23E-05
160	1,59E-05	1,59E-03	**580**	1,22E-06	1,22E-04	**1000**	4,10E-07	4,10E-05	**1420**	2,04E-07	2,04E-05	**1840**	1,21E-07	1,21E-05
170	1,41E-05	1,41E-03	**590**	1,18E-06	1,18E-04	**1010**	4,02E-07	4,02E-05	**1430**	2,01E-07	2,01E-05	**1850**	1,20E-07	1,20E-05
180	1,26E-05	1,26E-03	**600**	1,14E-06	1,14E-04	**1020**	3,95E-07	3,94E-05	**1440**	1,98E-07	1,98E-05	**1860**	1,19E-07	1,19E-05
190	1,13E-05	1,13E-03	**610**	1,10E-06	1,10E-04	**1030**	3,87E-07	3,87E-05	**1450**	1,95E-07	1,95E-05	**1870**	1,18E-07	1,18E-05
200	1,02E-05	1,02E-03	**620**	1,07E-06	1,07E-04	**1040**	3,80E-07	3,79E-05	**1460**	1,93E-07	1,93E-05	**1880**	1,16E-07	1,16E-05
210	9,24E-06	9,24E-04	**630**	1,03E-02	1,03E+00	**1050**	3,72E-07	3,72E-05	**1470**	1,90E-07	1,90E-05	**1890**	1,15E-07	1,15E-05
220	8,42E-06	8,42E-04	**640**	1,00E-06	1,00E-04	**1060**	3,65E-07	3,65E-05	**1480**	1,88E-07	1,87E-05	**1900**	1,14E-07	1,14E-05
230	7,71E-06	7,71E-04	**650**	9,70E-07	9,70E-05	**1070**	3,59E-07	3,59E-05	**1490**	1,85E-07	1,85E-05	**1910**	1,13E-07	1,13E-05
240	7,08E-06	7,08E-04	**660**	9,41E-07	9,41E-05	**1080**	3,52E-07	3,52E-05	**1500**	1,83E-07	1,83E-05	**1920**	1,12E-07	1,12E-05
250	6,53E-06	6,53E-04	**670**	9,13E-07	9,13E-05	**1090**	3,46E-07	3,46E-05	**1510**	1,80E-07	1,80E-05	**1930**	1,10E-07	1,10E-05
260	6,04E-06	6,04E-04	**680**	8,87E-07	8,87E-05	**1100**	3,39E-07	3,39E-05	**1520**	1,78E-07	1,78E-05	**1940**	1,09E-07	1,09E-05
270	5,60E-06	5,60E-04	**690**	8,61E-07	8,61E-05	**1110**	3,33E-07	3,33E-05	**1530**	1,75E-07	1,75E-05	**1950**	1,08E-07	1,08E-05
280	5,21E-06	5,21E-04	**700**	8,37E-07	8,37E-05	**1120**	3,27E-07	3,27E-05	**1540**	1,73E-07	1,73E-05	**1960**	1,07E-07	1,07E-05
290	4,86E-06	4,86E-04	**710**	8,14E-07	8,13E-05	**1130**	3,22E-07	3,22E-05	**1550**	1,71E-07	1,71E-05	**1970**	1,06E-07	1,06E-05
300	4,54E-06	4,54E-04	**720**	7,91E-07	7,91E-05	**1140**	3,16E-07	3,16E-05	**1560**	1,69E-07	1,69E-05	**1980**	1,05E-07	1,05E-05
310	4,25E-06	4,25E-04	**730**	7,70E-07	7,70E-05	**1150**	3,10E-07	3,10E-05	**1570**	1,67E-07	1,67E-05	**1990**	1,04E-07	1,04E-05
320	3,99E-06	3,99E-04	**740**	7,49E-07	7,49E-05	**1160**	3,05E-07	3,05E-05	**1580**	1,65E-07	1,65E-05	**2000**	1,03E-07	1,03E-05
330	3,75E-06	3,75E-04	**750**	7,29E-07	7,29E-05	**1170**	3,00E-07	3,00E-05	**1590**	1,62E-07	1,62E-05	**2010**	1,02E-07	1,02E-05
340	3,54E-06	3,54E-04	**760**	7,10E-07	7,10E-05	**1180**	2,95E-07	2,95E-05	**1600**	1,61E-07	1,60E-05	**2020**	1,01E-07	1,01E-05
350	3,34E-06	3,34E-04	**770**	6,92E-07	6,92E-05	**1190**	2,90E-07	2,90E-05	**1610**	1,59E-07	1,59E-05	**2030**	9,97E-08	9,97E-06
360	3,16E-06	3,16E-04	**780**	6,74E-07	6,74E-05	**1200**	2,85E-07	2,85E-05	**1620**	1,57E-07	1,56E-05			
370	2,99E-06	2,99E-04	**790**	6,57E-07	6,57E-05	**1210**	2,80E-07	2,80E-05	**1630**	1,55E-07	1,55E-05			
380	2,83E-06	2,83E-04	**800**	6,41E-07	6,41E-05	**1220**	2,76E-07	2,76E-05	**1640**	1,53E-07	1,53E-05			
390	2,69E-06	2,69E-04	**810**	6,25E-07	6,25E-05	**1230**	2,71E-07	2,71E-05	**1650**	1,51E-07	1,51E-05	Πέρας αλγορίθμου		
400	2,56E-06	2,56E-04	**820**	6,10E-07	6,10E-05	**1240**	2,67E-07	2,67E-05	**1660**	1,49E-07	1,49E-05			
410	2,43E-06	2,43E-04	**830**	5,96E-07	5,96E-05	**1250**	2,63E-07	2,63E-05	**1670**	1,47E-07	1,47E-05			
420	2,32E-06	2,32E-04	**840**	5,81E-07	5,81E-05	**1260**	2,59E-07	2,59E-05	**1680**	1,46E-07	1,45E-05			

Πίνακας 3.3.1.Α. Τιμές της νόρμα άπειρο για απόλυτο και ποσοστιαίο σφάλμα με διαφορετικό πλήθος σημείων.



Γράφημα 3.3.1.Α. Τιμές της μεγίστης νόρμας σε λογαριθμική κλίμακα για απόλυτο και ποσοστιαίο σφάλμα.

4. Η μέθοδος Galerkin

4.1. Η μέθοδος Galerkin στη γενική περίπτωση

Τα βήματα που ακολουθούμε για την εύρεση της προσέγγισης με την μέθοδο Galerkin, είναι τα ακόλουθα :

1ο βήμα: Κατασκευάζουμε την ασθενή μορφή πολλαπλασιάζοντας την εξίσωση με μια δοκιμαστική συνάρτηση και ολοκληρώνοντας στο αντίστοιχο πεδίο ορισμού.

2ο βήμα: Παράγουμε ενα πλέγμα, π.χ. $x_0,...,x_M$ και ορίζουμε:

- Ως κόμβο ή κομβικό σημείο να είναι τα $x_i, \;\; i=0,...,M$.
- Ως εκλέπτυνση του πλέγματος (βήμα) να είναι η $h = \max\limits_{0\le i\le M-1}\{h_i\}$ με $h_i = x_{i+1} - x_i, \;\; i=0,...,M-1$.
- Ως στοιχείο e_i του πλέγματος να είναι το διάστημα $e_i = [x_i, x_{i+1}], \;\; i=0,...,M-1$.

Στόχος μας είναι να κατασκευάσουμε έναν πεπερασμένων διαστάσεων χώρο V_H, έτσι ώστε να μπορούμε να εκφράσουμε κάθε σημείο του ως γραμμικό συνδυασμό των συναρτήσεων βάσης $\varphi_1, \varphi_2, ..., \varphi_{M-1}$ και ο απλούστερος που μπορούμε να παράξουμε, είναι ο χώρος των τμηματικά συνεχών γραμμικών συναρτήσεων :

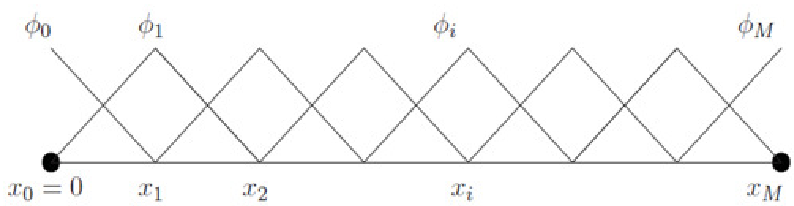
$$V_H = \left\{ u_h(x), \;\; u_h(x) = \sum_{j=1}^{M-1} \alpha_j \varphi_j \right\}$$

3ο βήμα: Θέτουμε οι συναρτήσεων βάσης είναι B splines 1ου βαθμού οι οποίες λαμβάνουν ως τιμή την μονάδα σε ένα σημείο και μηδέν αλλού, συνεπώς ισχύει ότι :

$$\begin{aligned}
&\varphi_1(x_1) = 1, \;\; \varphi_1(x_j) = 0, \;\; j = 0,2,3,...,M-1 \\
&\varphi_2(x_2) = 1, \;\; \varphi_2(x_j) = 0, \;\; j = 0,1,3,...,M-1 \\
&\vdots \\
&\varphi_i(x_3) = 1, \;\; \varphi_i(x_j) = 0, \;\; j = 0,...,i-1,i+1,...,M-1 \\
&\vdots \\
&\varphi_{M-1}(x_{M-1}) = 1, \;\; \varphi_{M-1}(x_j) = 0, \;\; j = 0,...,M-2
\end{aligned}$$

Από τα παραπάνω προκύπτει ότι η αναλυτική μορφή της συνάρτησης, είναι η εξής:

$$\varphi_i(x) = \begin{cases} 0 & \text{, εάν } x < x_{i-1} \\ \dfrac{x - x_{i-1}}{h_i} & \text{, εάν } x_{i-1} \le x < x_i \\ \dfrac{x_{i+1} - x}{h_{i+1}} & \text{, εάν } x_i \le x < x_{i+1} \\ 0 & \text{, αλλιώς} \end{cases}$$



Στο σημείο αυτό θέλουμε να σημειώσουμε ότι όταν χρησιμοποιούμε ένα πλέγμα Μ+1 στοιχείων $x_0,...,x_M$ για τον προσδιορισμό του V_H χρειάζεται να προσδιορίσουμε Μ-1 διανύσματα βάσης ή ισοδύναμα τις $\varphi_1,\varphi_2,...,\varphi_{M-1}$. Το σκεπτικό είναι ότι η γραμμική συνάρτηση $l(x)$ με κανόνα αντιστοίχησης :

$$l(x) = l(x_i)\cdot\left(\frac{x - x_{i+1}}{x_i - x_{i+1}}\right) + l(x_{i+1})\cdot\left(\frac{x - x_i}{x_{i+1} - x_i}\right)$$

, στο διάστημα $[x_i, x_{i+1}]$ προσδιορίζεται κατά μοναδικό τρόπο από τις τιμές x_i και x_{i+1} διότι υπάρχουν Μ-1 τέτοιες τιμές για τους κόμβους και περαιτέρω ισχύει ότι $l(x_0) = l(x_M) = 0$. Συνεπώς δοθέντος ενός διανύσματος $[l(x_1), l(x_2), ..., l(x_{M-1})]^T \in R^{M-1}$, μπορούμε να κατασκευάσουμε μια συνάρτηση $u_h(x) \in V_h$ εάν λάβουμε ως $u_h(x) = l(x_i),\ i = 1,...,M-1$ και δοθείσας μιας συνάρτησης $u_h(x) \in V_h$, μπορούμε να κατασκευάσουμε ένα διάνυσμα της μορφής $[v(x_1), v(x_2), ..., v(x_{M-1})]^T \in R^{M-1}$. Κατ' αυτόν τον τρόπο κατασκευάζουμε μια ένα προς ένα αντιστοίχιση από τον χώρο V_h στον χώρο R^{M-1} ή ισοδύναμα ο χώρος V_h είναι πεπερασμένος με $\dim(V_h) = M-1$.

4ο βήμα: Για κάθε ένα στοιχείο κατασκευάζουμε τους πίνακες τοπικής δυσκαμψίας και τοπικού φορτίου, η μορφή των οποίων εξαρτάται από την εξίσωση που αντιμετωπίζουμε και αντιπροσωπεύουν την συνεισφορά κάθε στοιχείου στους πίνακες A και F του συστήματος AU=F .

5ο βήμα: Κατασκευάζουμε τον καθολικό πίνακα δυσκαμψίας A και το καθολικό διάνυσμα φορτίου F αντίστοιχα, αθροίζοντας στοιχειό προς στοιχείο τους αντίστοιχους local stiffness και local load vector πίνακες.

6ο βήμα: Λύνουμε το σύστημα AU=F και ο πίνακας U που προκύπτει είναι η προσέγγιση της λύσης.

7ο βήμα: Εκτιμούμε το σφάλμα.

4.2. Εφαρμογή στη μια διάσταση

Παρακάτω αναλύουμε την εφαρμογή της μεθοδολογίας που αναπτύξαμε στην προηγούμενη παράγραφο στην εξίσωση Poisson στη μια διάσταση :

$$-u''(x) = f(x),\ \ 0 < x < 1,\ \ u(0) = 0,\ \ u(1) = 0$$

, με $\Omega \subset R$ να είναι το πεδίο ορισμού στο επίπεδο (x,y) με σύνορο $\partial\Omega$.

1ο βήμα: Παράγουμε την ασθενή μορφή :

Δεν είναι δυνατή η εμφάνιση της εικόνας. Ίσως να μην επαρκεί η μνήμη του υπολογιστή για το άνοιγμα της εικόνας ή ίσως η εικόνα να έχει καταστραφεί. Επανεκκινήστε τον υπολογιστή και ανοίξτε ξανά το αρχείο. Αν εμφανίζεται ακόμα το κόκκινο x, ίσως να πρέπει να διαγράψετε την εικόνα και να την εισαγάγετε ξανά.

Από την παραπάνω ασθενή μορφή, εάν ορίσουμε τις συναρτήσεις :

$$\alpha(u,v) = \int_{\Omega} \left(u' \cdot v'\right) dx, \quad l(v) = \int_{\Omega} \left(f \cdot v\right) dx$$

, τότε προκύπτει ότι η $\alpha(u,v)$ είναι διγραμμική μορφή, η $l(v)$ είναι γραμμική μορφή και περαιτέρω εκπληρούν τα κριτήρια του θεωρήματος Lax – Milgram [ ]. Εφόσον εμείς επιλέγουμε την μορφή της δοκιμαστικής συνάρτησης v εάν θέσουμε v=u ή v=f , λαμβάνουμε αντίστοιχα :

$$\int_0^1 u'v'dx = \int_0^1 \left(u'\right)^2 dx \quad , \quad \int_0^1 fvdx = \int_0^1 f^2 dx .$$

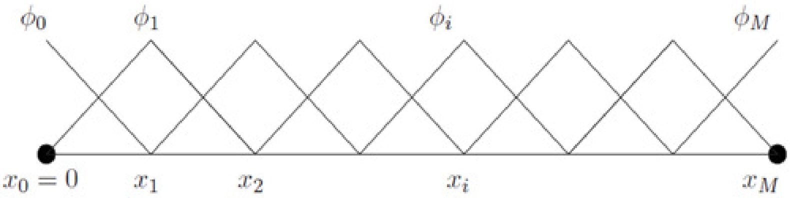
2ο βήμα: Παράγουμε ενα πλέγμα για να σχηματίσουμε τα διαστήματα $[x_{i-1}, x_i],\ i = 1,...,M$.

3ο βήμα: Θεωρούμε ότι το σύνολο των συναρτήσεων βάσης είναι οι συναρτήσεις hat :

$$\varphi_i(x) = \begin{cases} 0 & ,\text{εάν } x < x_{i-1} \\ \dfrac{x - x_{i-1}}{h_i} & ,\text{εάν } x_{i-1} \le x < x_i \\ \dfrac{x_{i+1} - x}{h_{i+1}} & ,\text{εάν } x_i \le x < x_{i+1} \\ 0 & ,\text{αλλιώς} \end{cases}$$

, π.χ. Για την συνάρτηση βάσης $\varphi_1(x)$ ισχύει ότι :

$$\varphi_1(x) = \begin{cases} 0 & ,\text{εάν } x < x_0 \\ \dfrac{x - x_0}{h_1} & ,\text{εάν } x_0 \le x < x_1 \\ \dfrac{x_2 - x}{h_2} & ,\text{εάν } x_1 \le x < x_2 \\ 0 & ,\text{αλλιώς} \end{cases}$$



Η λύση $u_h(x)$ θα είναι γραμμικός συνδυασμός των συναρτήσεων βάσης, συνεπώς θα είναι της μορφής $u_h(x)=\sum_{i=1}^{M-1} c_j\varphi_j(x)$, όπου οι συντελεστές c_j είναι άγνωστοι και για τον υπολογισμό των συντελεστών παράγουμε ένα γραμμικό σύστημα από την ασθενή μορφή ότι $\int_0^1 (u'v')dx=\int_0^1 (fv)dx$, επιλέγοντας κάθε φορά ως $v(x)$ να είναι αντίστοιχα μια από τις συναρτήσεις $\varphi_1,\ldots,\varphi_{M-1}$:

$$
\begin{aligned}
&v=\varphi_1,\quad c_1\left(\int_0^1 \varphi_1'\varphi_1' dx\right)+c_2\left(\int_0^1 \varphi_1'\varphi_2' dx\right)+\ldots+c_{n-1}\left(\int_0^1 \varphi_1'\varphi_{n-1}' dx\right)=\left(\int_0^1 f\varphi_1 dx\right)\\
&v=\varphi_2,\quad c_1\left(\int_0^1 \varphi_2'\varphi_1' dx\right)+c_2\left(\int_0^1 \varphi_2'\varphi_2' dx\right)+\ldots+c_{n-1}\left(\int_0^1 \varphi_2'\varphi_{n-1}' dx\right)=\left(\int_0^1 f\varphi_2 dx\right)\\
&\qquad\qquad\qquad\vdots\\
&v=\varphi_i,\quad c_1\left(\int_0^1 \varphi_i'\varphi_1' dx\right)+c_2\left(\int_0^1 \varphi_i'\varphi_2' dx\right)+\ldots+c_{n-1}\left(\int_0^1 \varphi_i'\varphi_{n-1}' dx\right)=\left(\int_0^1 f\varphi_i dx\right)\\
&\vdots\\
&v=\varphi_{M-1},\quad c_1\left(\int_0^1 \varphi_{M-1}'\varphi_1' dx\right)+c_2\left(\int_0^1 \varphi_{n-1}'\varphi_2' dx\right)+\ldots+c_{M-1}\left(\int_0^1 \varphi_{M-1}'\varphi_{M-1}' dx\right)=\left(\int_0^1 f\varphi_{M-1} dx\right)
\end{aligned}
$$

Εάν αναδιατυπώσουμε το παραπάνω σύστημα σε μορφή πινάκων $AU=F$, προκύπτει ότι :

$$
\begin{bmatrix}
\alpha(\varphi_1,\varphi_1) & \alpha(\varphi_1,\varphi_2) & \ldots & \alpha(\varphi_1,\varphi_{M-1})\\
\alpha(\varphi_2,\varphi_1) & \alpha(\varphi_2,\varphi_2) & \ldots & \alpha(\varphi_2,\varphi_{M-1})\\
\vdots & \vdots & \vdots & \vdots\\
\alpha(\varphi_{M-1},\varphi_1) & \alpha(\varphi_{M-1},\varphi_2) & \ldots & \alpha(\varphi_{M-1},\varphi_{M-1})
\end{bmatrix}
\begin{bmatrix} c_1\\ c_2\\ \vdots\\ c_{M-1}\end{bmatrix}
=
\begin{bmatrix} (f,\varphi_1)\\ (f,\varphi_2)\\ \vdots\\ (f,\varphi_{M-1})\end{bmatrix}
$$

, με $\alpha(\varphi_i,\varphi_j)=\left(\int_0^1 \varphi_i'\varphi_j' dx\right)$, $(f,\varphi_i)=\left(\int_0^1 f\varphi_i dx\right)$ ή ισοδύναμα το αρχικό πρόβλημα ανάγεται στην λύση του συστήματος $AU=F$.

4^{o} βήμα: Για τους πίνακες A και F η μεθοδολογία που ακολουθείται είναι η προσέγγιση της δόμησης στοιχείο προς στοιχείο, με κάθε ένα στοιχείο να είναι της μορφής:

$$\underbrace{[x_0,x_1]}_{e_1},\underbrace{[x_1,x_2]}_{e_2},\ldots,\underbrace{[x_{i-1},x_i]}_{e_i},\ldots,\underbrace{[x_{M-1},x_M]}_{e_M}$$

Η κεντρική ιδέα είναι να σπάσουμε την ολοκλήρωση στοιχείο προς στοιχείο, όπου για κάθε ολοκληρώσιμη συνάρτηση g(x), θα έχουμε ότι:

$$\int_\Omega g(x)dx=\int_0^1 g(x)dx=\sum_{k=1}^{M}\int_{x_{k-1}}^{x_k} g(x)dx=\sum_{k=1}^{M}\int_{\Omega_k} g(x)dx$$

$$A=\begin{bmatrix}\int_0^1(\varphi_1')^2dx & \int_0^1(\varphi_1')(\varphi_2')dx & \dots & \int_0^1(\varphi_1')(\varphi_{M-1}')dx\\ \int_0^1(\varphi_2')(\varphi_1')dx & \int_0^1(\varphi_2')^2dx & \dots & \int_0^1(\varphi_2')(\varphi_{M-1}')dx\\ \vdots & \vdots & \vdots & \vdots\\ \int_0^1(\varphi_{M-1}')(\varphi_1')dx & \int_0^1(\varphi_{M-1}')(\varphi_2')dx & \dots & \int_0^1(\varphi_{M-1}')^2dx\end{bmatrix}\Leftrightarrow$$

$$A=\begin{bmatrix}\int_{x_0}^{x_1}(\varphi_1')^2dx & \int_{x_0}^{x_1}(\varphi_1')(\varphi_2')dx & \dots & \int_{x_0}^{x_1}(\varphi_1')(\varphi_{M-1}')dx\\ \int_{x_0}^{x_1}(\varphi_2')(\varphi_1')dx & \int_{x_0}^{x_1}(\varphi_2')^2dx & \dots & \int_{x_0}^{x_1}(\varphi_2')(\varphi_{M-1}')dx\\ \vdots & \vdots & \vdots & \vdots\\ \int_{x_0}^{x_1}(\varphi_{M-1}')(\varphi_1')dx & \int_{x_0}^{x_1}(\varphi_{M-1}')(\varphi_2')dx & \dots & \int_{x_0}^{x_1}(\varphi_{M-1}')^2dx\end{bmatrix}+\begin{bmatrix}\int_{x_1}^{x_2}(\varphi_1')^2dx & \int_{x_1}^{x_2}(\varphi_1')(\varphi_2')dx & \dots & \int_{x_1}^{x_2}(\varphi_1')(\varphi_{M-1}')dx\\ \int_{x_1}^{x_2}(\varphi_2')(\varphi_1')dx & \int_{x_1}^{x_2}(\varphi_2')^2dx & \dots & \int_{x_1}^{x_2}(\varphi_2')(\varphi_{M-1}')dx\\ \vdots & \vdots & \vdots & \vdots\\ \int_{x_1}^{x_2}(\varphi_{M-1}')(\varphi_1')dx & \int_{x_1}^{x_2}(\varphi_{M-1}')(\varphi_2')dx & \dots & \int_{x_1}^{x_2}(\varphi_{M-1}')^2dx\end{bmatrix}+\dots+\begin{bmatrix}\int_{x_{M-1}}^{x_M}(\varphi_1')^2dx & \int_{x_{M-1}}^{x_M}(\varphi_1')(\varphi_2')dx & \dots & \int_{x_{M-1}}^{x_M}(\varphi_1')(\varphi_{M-1}')dx\\ \int_{x_{M-1}}^{x_M}(\varphi_2')(\varphi_1')dx & \int_{x_{M-1}}^{x_M}(\varphi_2')^2dx & \dots & \int_{x_{M-1}}^{x_M}(\varphi_2')(\varphi_{M-1}')dx\\ \vdots & \vdots & \vdots & \vdots\\ \int_{x_{M-1}}^{x_M}(\varphi_{M-1}')(\varphi_1')dx & \int_{x_{M-1}}^{x_M}(\varphi_{M-1}')(\varphi_2')dx & \dots & \int_{x_{M-1}}^{x_M}(\varphi_{M-1}')^2dx\end{bmatrix}$$

Εφόσον για τις hat functions νωρίζουμε ότι σε κάθε διάστημα υπάρχουν ακριβώς δυο μη μηδενικές συναρτήσεις βάσης, θα έχουμε ότι ο παραπάνω πίνακας λαμβάνει την εξής μορφή :

$$A=\underbrace{\begin{bmatrix}\int_{x_0}^{x_1}(\varphi_1')^2dx & 0 & \dots & 0\\ 0 & 0 & \dots & 0\\ \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & \dots & 0\end{bmatrix}}_{K^0}+\underbrace{\begin{bmatrix}\int_{x_1}^{x_2}(\varphi_1')^2dx & \int_{x_1}^{x_2}(\varphi_1')(\varphi_2')dx & \dots & 0\\ \int_{x_1}^{x_2}(\varphi_2')(\varphi_1')dx & \int_{x_1}^{x_2}(\varphi_2')^2dx & \dots & 0\\ \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & \dots & 0\end{bmatrix}}_{K^1}+\underbrace{\begin{bmatrix}0 & 0 & 0 & \dots & 0\\ 0 & \int_{x_2}^{x_3}(\varphi_2')^2dx & \int_{x_2}^{x_3}(\varphi_2')(\varphi_3')dx & \dots & 0\\ 0 & \int_{x_2}^{x_3}(\varphi_3')(\varphi_2')dx & \int_{x_2}^{x_3}(\varphi_3')^2dx & \dots & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & 0 & \dots & 0\end{bmatrix}}_{K^2}+\dots+\underbrace{\begin{bmatrix}0 & 0 & 0 & \dots & 0\\ 0 & 0 & 0 & \dots & 0\\ 0 & 0 & 0 & \dots & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & 0 & \dots & \int_{x_{M-1}}^{x_M}(\varphi_{M-1}')^2dx\end{bmatrix}}_{K^{M-1}}$$

Ως μη μηδενική συνεισφορά του κάθε στοιχείου $e=[x_i,x_{i+1}]$ ορίζουμε να είναι ο 2x2 πίνακας K_i^e, ο οποίος μπορεί να λάβει τις εξής τρείς μορφές:

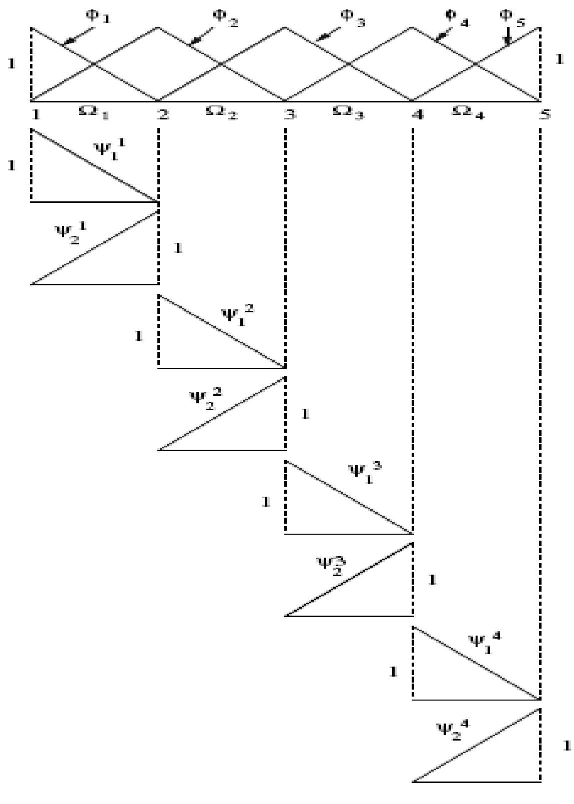
$$K_0^e=\begin{bmatrix}\int_{x_0}^{x_{i+1}}(\varphi_1^{'})^2dx & 0\\ 0 & 0\end{bmatrix},\quad K_{M-1}^e=\begin{bmatrix}0 & 0\\ 0 & \int_{x_i}^{x_{i+1}}(\varphi_{M-1}^{'})^2dx\end{bmatrix}$$

$$K_i^e=\begin{bmatrix}\int_{x_i}^{x_{i+1}}(\varphi_i^{'})^2dx & \int_{x_i}^{x_{i+1}}(\varphi_i^{'})(\varphi_{i+1}^{'})dx\\ \int_{x_i}^{x_{i+1}}(\varphi_{i+1}^{'})(\varphi_i^{'})dx & \int_{x_i}^{x_{i+1}}(\varphi_{i+1}^{'})^2dx\end{bmatrix},\quad i=1,...,M-2$$

Για τον υπολογισμό του τοπικού πίνακα δυσκαμψίας K_i^e χρησιμοποιούμε το γεγονός ότι στο στοιχείο $e=[x_i,x_{i+1}]$, η συνάρτηση :

$$\varphi_i(x)=\begin{cases}0 & ,\text{εάν } x<x_{i-1}\\ \dfrac{x-x_{i-1}}{h_i} & ,\text{εάν } x_{i-1}\le x<x_i\\ \dfrac{x_{i+1}-x}{h_{i+1}} & ,\text{εάν } x_i\le x<x_{i+1}\\ 0 & ,\text{αλλιώς}\end{cases}$$

, μπορεί να αναλυθεί στις μη μηδενικές συναρτήσεων $\psi_i^e(x)=\dfrac{x_{i+1}-x}{x_{i+1}-x_i}$ και $\psi_{i+1}^e(x)=\dfrac{x-x_i}{x_{i+1}-x_i}$ [7]:



, για τις παραγώγους των οποίων έχουμε ότι :

$$\left(\psi_i^e(x)\right)' = -\frac{1}{h_i}, \quad \left(\psi_{i+1}^e(x)\right)' = \frac{1}{h_i}$$

Συνεπώς για την συνεισφορά τους στα αντίστοιχα ολοκληρώματα, προκύπτει ότι :

$$\int_{x_i}^{x_{i+1}} (\varphi_i^{'})^2 dx = \int_{x_0}^{x_1} \left(\psi_i^{'}\right)^2 dx = \int_{x_0}^{x_1} \left(\frac{1}{h_i}\right)^2 dx = \frac{1}{h_i}$$

$$\int_{x_i}^{x_{i+1}} (\varphi_{i+1}^{'})^2 dx = \int_{x_0}^{x_1} \left(\psi_{i+1}^{'}\right)^2 dx = \int_{x_0}^{x_1} \left(\frac{1}{h_i}\right)^2 dx = \frac{1}{h_i}$$

$$\int_{x_i}^{x_{i+1}} (\varphi_i^{'})(\varphi_{i+1}^{'}) dx = \int_{x_0}^{x_1} \psi_i^{'}\varphi_{i+1}^{'} dx = \int_{x_0}^{x_1} -\frac{1}{h_i}^2 dx = -\frac{1}{h_i}, \quad \int_{x_i}^{x_{i+1}} (\varphi_{i+1}^{'})(\varphi_i^{'}) dx = \int_{x_0}^{x_1} \varphi_{i+1}^{'}\psi_i^{'} dx = \int_{x_0}^{x_1} -\frac{1}{h_i}^2 dx = -\frac{1}{h_i}$$

Συνοψίζοντας, για τον τοπικό πίνακα δυσκαμψίας K_i^e, θα έχουμε ότι :

$$K_i^e = \begin{bmatrix} \int_{x_i}^{x_{i+1}} (\varphi_i^{'})^2 dx & \int_{x_i}^{x_{i+1}} (\varphi_i^{'})(\varphi_{i+1}^{'}) dx \\ \int_{x_i}^{x_{i+1}} (\varphi_{i+1}^{'})(\varphi_i^{'}) dx & \int_{x_i}^{x_{i+1}} (\varphi_{i+1}^{'})^2 dx \end{bmatrix} = \begin{bmatrix} \frac{1}{h_i} & -\frac{1}{h_i} \\ -\frac{1}{h_i} & \frac{1}{h_i} \end{bmatrix}$$

Κατ' αναλογία το τοπικό διάνυσμα φορτίου F_i^e του κάθε στοιχείου ορίζεται ως εξής :

$$F_i^e = \begin{bmatrix} \int_{x_i}^{x_{i+1}} \left(f\varphi_i\right) dx \\ \int_{x_i}^{x_{i+1}} \left(f\varphi_{i+1}\right) dx \end{bmatrix}$$

, όπου η μορφή του εξαρτάται από την f.

5^{o} βήμα : Από το 4^{o} βήμα θα έχουμε ότι $A = K_0^e + K_1^e + \ldots + K_{M-1}^e$, όπου για τους K_i^e ισχύει :

$$K_0^e = \begin{bmatrix} \int_{x_0}^{x_1} (\varphi_1^{'})^2 dx & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{h_0} & 0 \\ 0 & 0 \end{bmatrix}, \quad K_{M-1}^e = \begin{bmatrix} 0 & 0 \\ 0 & \int_{x_i}^{x_{i+1}} (\varphi_{M-1}^{'})^2 dx \end{bmatrix}, = \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{h_{M-1}} \end{bmatrix},$$

$$K_i^e = \begin{bmatrix} \int_{x_i}^{x_{i+1}} (\varphi_i^{'})^2 dx & \int_{x_i}^{x_{i+1}} (\varphi_i^{'})(\varphi_{i+1}^{'}) dx \\ \int_{x_i}^{x_{i+1}} (\varphi_{i+1}^{'})(\varphi_i^{'}) dx & \int_{x_i}^{x_{i+1}} (\varphi_{i+1}^{'})^2 dx \end{bmatrix} = \begin{bmatrix} \frac{1}{h_i} & -\frac{1}{h_i} \\ -\frac{1}{h_i} & \frac{1}{h_i} \end{bmatrix}$$

, συνεπώς ο καθολικός πίνακας δυσκαμψίας A για την συγκεκριμένη εξίσωση προκύπτει αθροίζοντας τους παραπάνω πίνακες, οπότε λαμβάνουμε τον τριδιαγώνιο πίνακα A :

$$A=\begin{bmatrix} \frac{1}{h_0}+\frac{1}{h_1} & -\frac{1}{h_1} & & & & \\ -\frac{1}{h_1} & \frac{1}{h_1}+\frac{1}{h_2} & -\frac{1}{h_2} & & & \\ & -\frac{1}{h_2} & \frac{1}{h_2}+\frac{1}{h_3} & -\frac{1}{h_3} & & \\ & & \ddots & \ddots & \ddots & \\ & & & -\frac{1}{h_{M-3}} & \frac{1}{h_{M-3}}+\frac{1}{h_{M-2}} & -\frac{1}{h_{M-3}} \\ & & & & -\frac{1}{h_{M-2}} & \frac{1}{h_{M-2}}+\frac{1}{h_{M-1}} \end{bmatrix}$$

Γνωρίζουμε ότι για το καθολικό διάνυσμα φορτίου ισχύει ότι $F=F_0^e+F_1^e+...+F_{M-1}^e$ όπου για τους F_i^e ισχύει :

$$F_0^e=\begin{bmatrix} \int_{x_i}^{x_{i+1}}(f\varphi_{i+1})dx \\ 0 \end{bmatrix},\quad F_{M-1}^e=\begin{bmatrix} 0 \\ \int_{x_i}^{x_{i+1}}(f\varphi_{i+1})dx \end{bmatrix},\quad F_i^e=\begin{bmatrix} \int_{x_i}^{x_{i+1}}(f\varphi_i)dx \\ \int_{x_i}^{x_{i+1}}(f\varphi_{i+1})dx \end{bmatrix},\quad i=1,...,M-2$$

, συνεπώς το καθολικό διάνυσμα φορτίου για την συγκεκριμένη εξίσωση προκύπτει αθροίζοντας τους παραπάνω πίνακες, οπότε λαμβάνουμε τον τριδιαγώνιο πίνακα F :

$$F=\begin{bmatrix} \int_{x_0}^{x_1}(f\varphi_1)dx \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}+\begin{bmatrix} \int_{x_1}^{x_2}(f\varphi_1)dx \\ \int_{x_1}^{x_2}(f\varphi_2)dx \\ 0 \\ \vdots \\ 0 \end{bmatrix}+\begin{bmatrix} 0 \\ \int_{x_2}^{x_3}(f\varphi_3)dx \\ \int_{x_3}^{x_4}(f\varphi_3)dx \\ \vdots \\ 0 \end{bmatrix}+...+\begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ \int_{x_{M-1}}^{x_M}(f\varphi_{M-1})dx \end{bmatrix}$$

, ο οποίος υπολογίζεται με οποιαδήποτε μέθοδο αριθμητικής ανάλυσης, διότι έχουμε ορισμένα ολοκληρώματα και η συνάρτηση f είναι γνωστή για το συγκεκριμένο πρόβλημα.

6^o βήμα: Βρίσκουμε την προσέγγιση της λύσης, λύνοντας το σύστημα που προκύπτει για την συγκεκριμένη συνάρτηση.

7^o βήμα: Εκτιμούμε το σφάλμα.

Σύνδεση των μορφών Ritz - Galerkin

Για το πρόβλημα:

$$\begin{aligned} & u''(x) = f(x),\ \ 0 < x < 1, \\ & u(0) = 0,\ \ u(1) = 0 \end{aligned}$$

, έχουμε ότι η μορφή ελαχιστοποίησης θα είναι η εξής:

$$\min_{v \in H^1(0,1)} F(v),\quad F(v) = \frac{1}{2}\int_0^1 \left(u_x\right)^2 dx - \int_0^1 \left(fv\right) dx$$

Κατ' αναλογία με την παρουσίαση της μεθοδολογίας, η λύση μας θα είναι μια προσέγγιση της μορφής $u_h(x) = \sum_{j=1}^{M-1} \alpha_j \varphi_j(x_i)$ και εν προκειμένω θα έχουμε ότι:

$$F\left(u_h(x)\right) = \frac{1}{2}\int_0^1 \left(\sum_{j=1}^{M-1} \alpha_j \varphi_j(x_i)\right)^2 dx - \int_0^1 f\left(\sum_{j=1}^{M-1} \alpha_j \varphi_j(x)\right) dx$$

Η παραπάνω συνάρτηση πολλών μεταβλητών, μπορεί να αναδιατυπωθεί ως εξής :

$$F(u_h) = F(\alpha_1, \alpha_2, ..., \alpha_{M-1})$$

Για το ολικό ελάχιστο χρησιμοποιούμε το κριτήριο της πρώτης παραγώγου :

$$\frac{\partial F}{\partial \alpha_1} = 0, ..., \frac{\partial F}{\partial \alpha_i} = 0, ..., \frac{\partial F}{\partial \alpha_{M-1}} = 0,$$

, συνεπώς προκύπτει ότι :

$$\begin{aligned} & \frac{\partial F}{\partial \alpha_1} = \int_0^1 \left(\sum_{j=1}^{M-1} \alpha_j \varphi_j'(x_i)\right) \varphi_1' dx - \int_0^1 f\varphi_1 dx = 0 \\ & \vdots \\ & \frac{\partial F}{\partial \alpha_{M-1}} = \int_0^1 \left(\sum_{j=1}^{M-1} \alpha_j \varphi_{M-1}'(x_i)\right) \varphi_{M-1}' dx - \int_0^1 f\varphi_{M-1} dx, \end{aligned}$$

ή ισοδύναμα εν γένει για την παράγωγο ισχύει ότι :

$$\frac{\partial F}{\partial \alpha_i} = \int_0^1 \left(\sum_{j=1}^{M-1} \alpha_j \varphi_j'(x_i)\right) \varphi_i' dx - \int_0^1 f\varphi_i dx = 0,\quad i = 1, ..., M-1$$

Εάν αλλάξουμε την ολοκλήρωση με την άθροιση θα έχουμε ότι πρέπει να λύσουμε το σύστημα :

$$\sum_{j=1}^{M-1} \left(\int_0^1 \varphi_j' \varphi_i' dx\right) \alpha_j = \int_0^1 f\varphi_i dx,\quad i = 1, ..., M-1$$

Για την μέθοδο Galerkin θα πρέπει να παράξουμε την ασθενή μορφή, συνεπώς πολλαπλασιάζουμε τα μέλη της εξίσωσης με μια δοκιμαστική συνάρτηση και ολοκληρώνουμε στο πεδίο ορισμού και λαμβάνουμε :

$$\int_0^1 \left(u'v\right)dx = \int_0^1 \left(fv\right)dx \Leftrightarrow \int_0^1 \left(\sum_{j=1}^{M-1} \alpha_j \varphi_j'(x_i)\right)\varphi_i' dx = \int_0^1 f\varphi_i dx, \quad i = 1,\dots,M-1$$

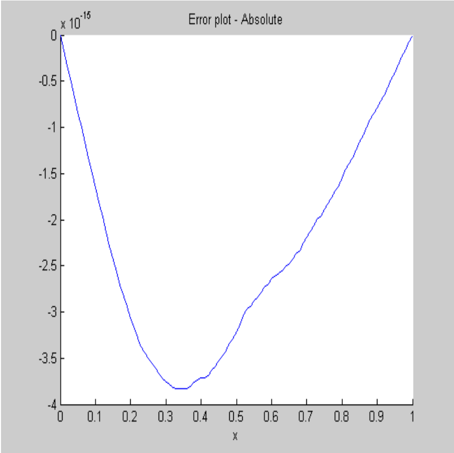
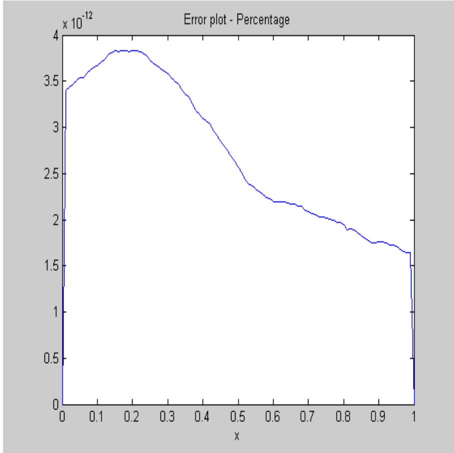
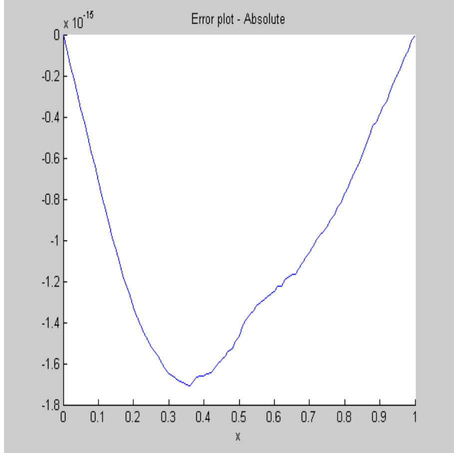
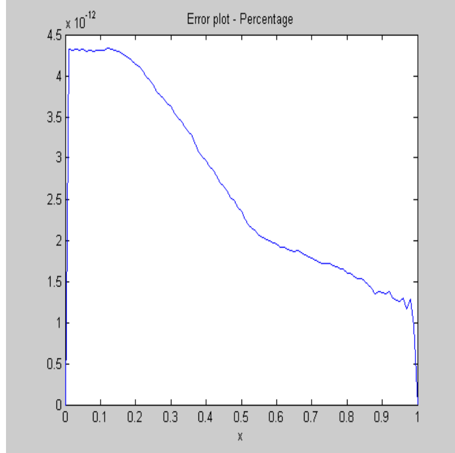
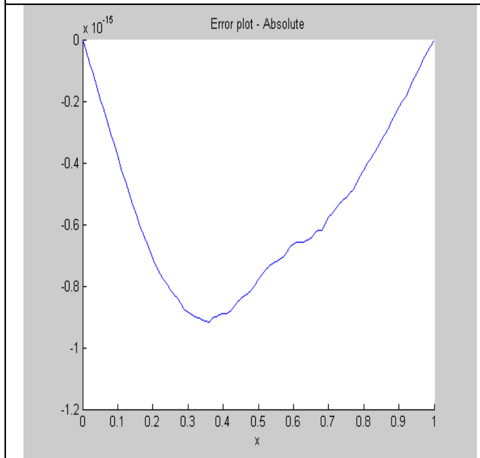
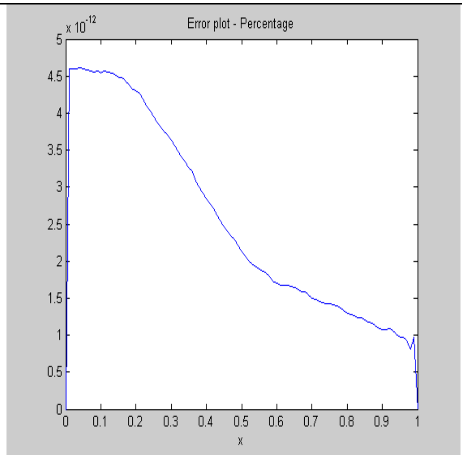
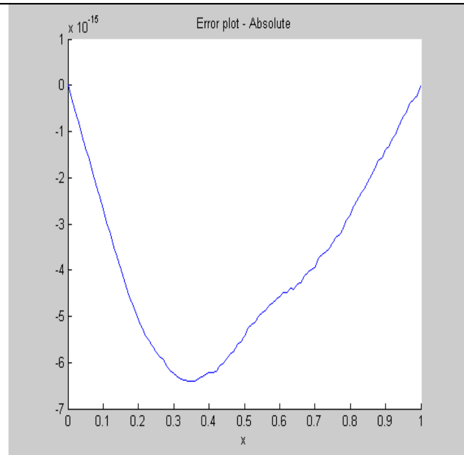
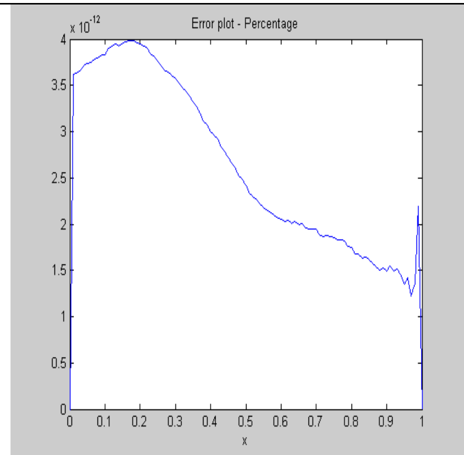
, συνεπώς το σύστημα που προκύπτει είναι το ίδιο.

4.2.1. Πειραματικά αποτελέσματα για την μέθοδο Galerking στη μια διάσταση

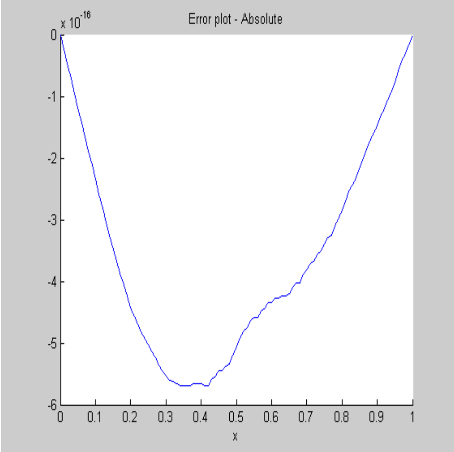
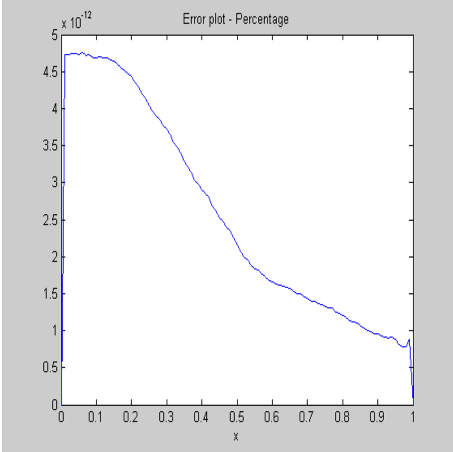
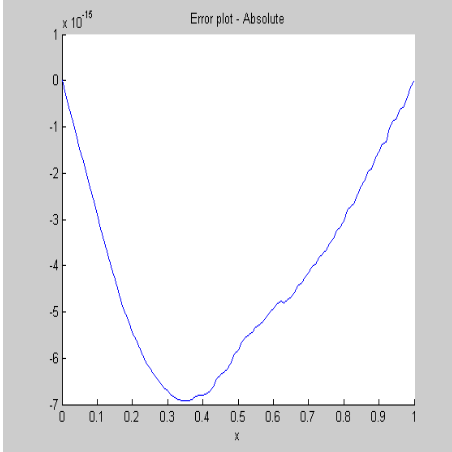
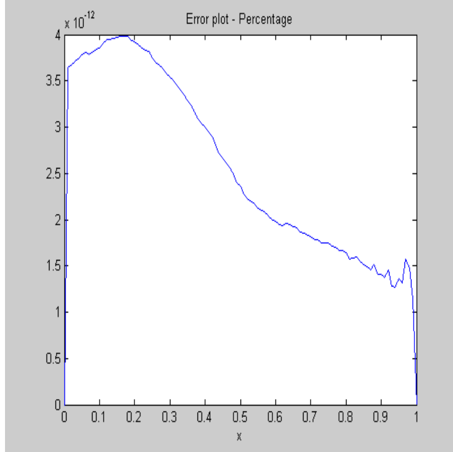
Για το μοντέλο $-u''(x) = f(x),\ 0 < x < 1,\ u(0) = 0,\ u(1) = 0$, εξετάσαμε τις εξής περιπτώσεις :

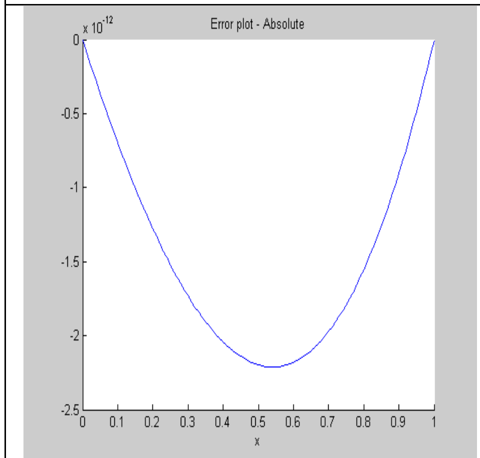
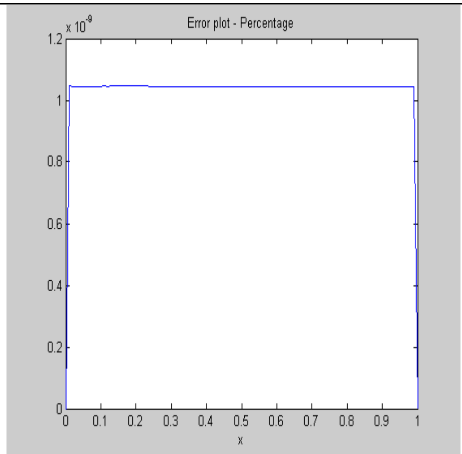
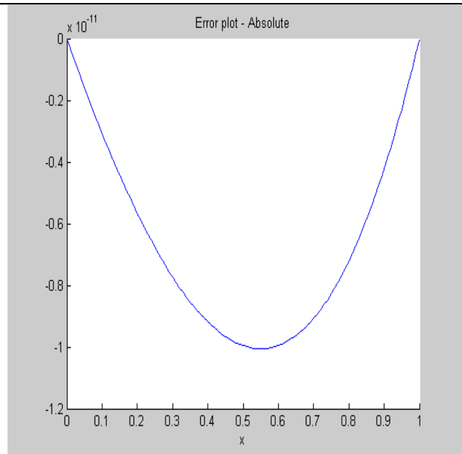
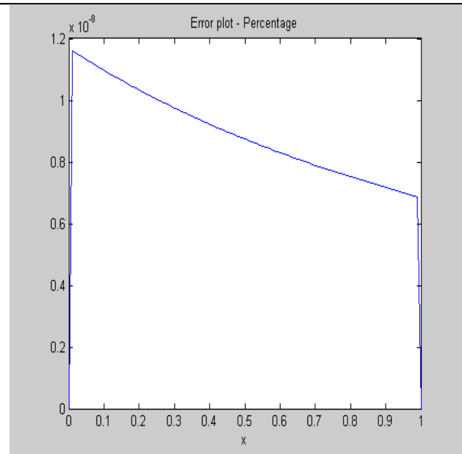
Τύπος συνάρτησης $f(x)$	**Αναλυτική μορφή λύσης**
$f_1(x) = 1$	$u(x) = -\frac{1}{2}\cdot x\cdot(x-1)$
$f_2(x) = x$	$u(x) = \frac{1}{6}\cdot(x - x^3)$
$f_3(x) = x^2$	$u(x) = \frac{1}{12}\cdot(x - x^4)$
$f_4(x) = x^2 + x + 1$	$u(x) = -\frac{1}{12}\cdot x\cdot\left(x^3 + 2x^2 + 6x - 9\right)$
$f_5(x) = x^3$	$u(x) = \frac{1}{20}\cdot\left(x - x^5\right)$
$f_6(x) = x^3 + x^2 + x + 1$	$u(x) = -\frac{1}{60}\cdot x\cdot\left(3x^4 + 5x^3 + 10x^2 + 30x - 48\right)$
$f_7(x) = e^x$	$u(x) = ex - x - e^x + 1$
$f_8(x) = e^x\cdot x$	$u(x) = -e^x(x-2) + 2(x-1) - ex$
$f_9(x) = e^x\cdot x^2$	$u(x) = -e^x(x^2 - 4x + 6) + 3ex - 6x + 6$
$f_{10}(x) = e^x\cdot(x^2 + x + 1)$	$u(x) = -e^x(x^2 - 3x + 5) + 3ex - 5x + 5$
$f_{11}(x) = e^x\cdot(x^3)$	$u(x) = -e^x(x^3 - 6x^2 + 18x - 24) + 24(x-1) - 11ex$
$f_{12}(x) = e^x\cdot(x^3 + x^2 + x + 1)$	$u(x) = -e^x(x^3 - 5x^2 + 15x - 19) + 19(x-1) - 8ex$
$f_{13}(x) = \eta\mu(x)$	$u(x) = \eta\mu(x) - x\eta\mu(1)$
$f_{14}(x) = \eta\mu(x)\cdot e^x$	$u(x) = -\frac{1}{2}\left(x - e\cdot x\cdot\cos(1) + e^x\cdot\cos(x) - 1\right)$

Πίνακας 4.2.1.Α. Δεκατέσσερις υλοποιήσεις της $f(x)$ και η αντίστοιχη αναλυτική λύση του παραγόμενου συστήματος.

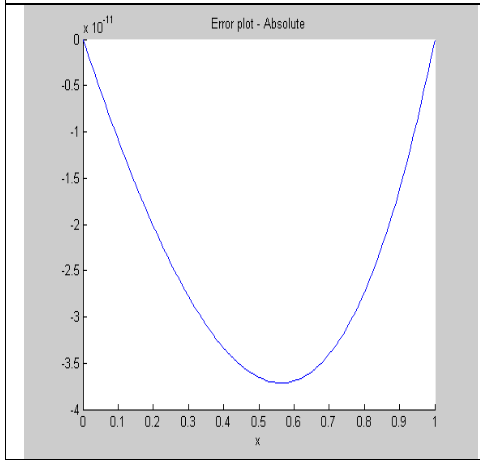
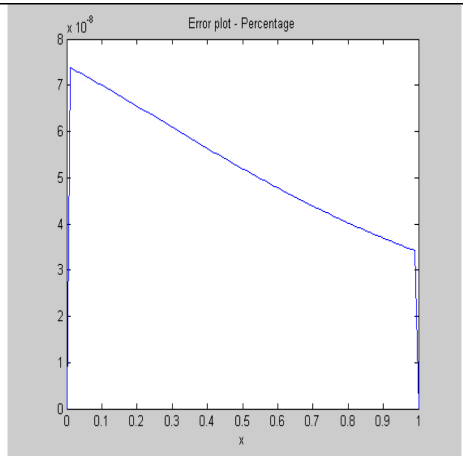
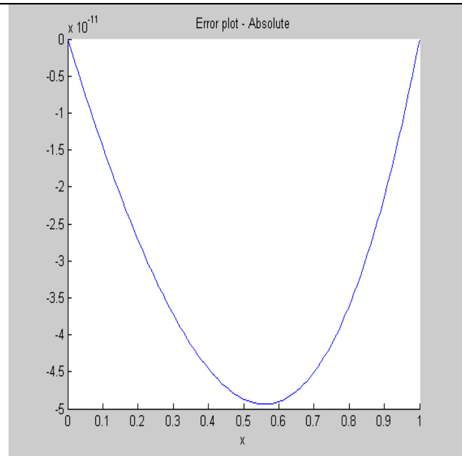
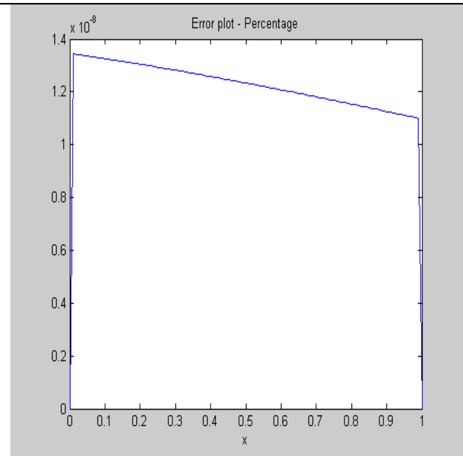
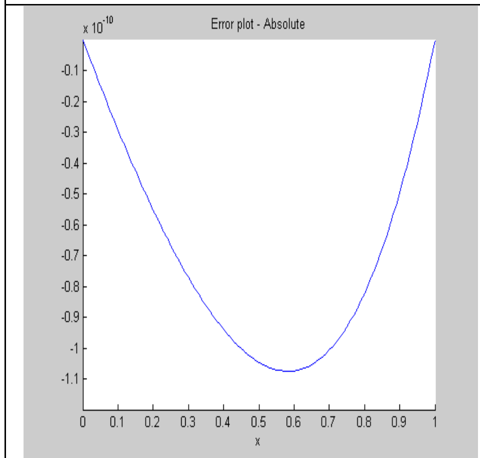
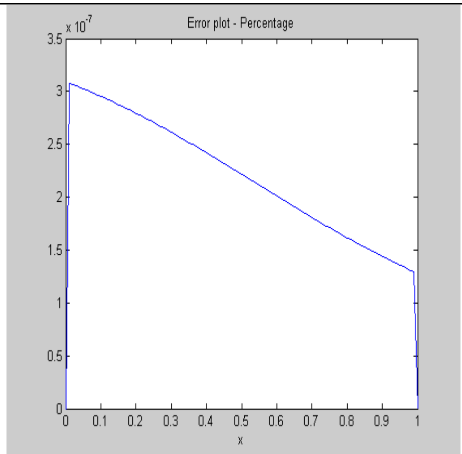
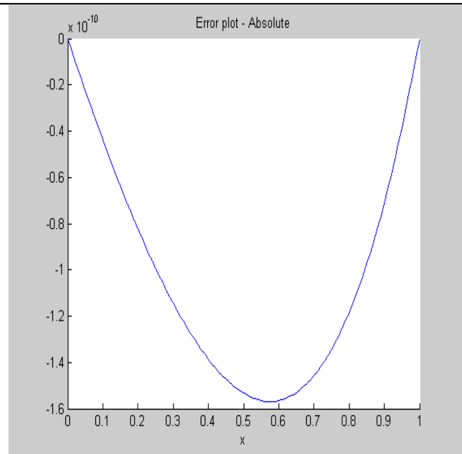
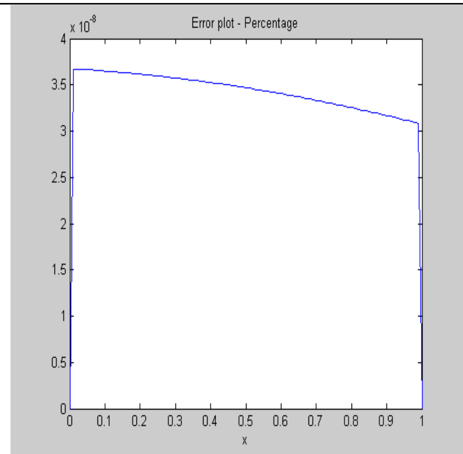
<table>
<tr><td colspan="2">$f_1(x)=1,\ u(x)=-\frac{1}{2}\cdot x\cdot(x-1)$</td><td colspan="2">$f_2(x)=x\ ,u(x)=\frac{1}{6}\cdot(x-x^3)$</td></tr>
<tr><td>Απόλυτο σφάλμα</td><td>Ποσοστιαίο σφάλμα</td><td>Απόλυτο σφάλμα</td><td>Ποσοστιαίο σφάλμα</td></tr>
<tr><td>3.8303e-015</td><td>3.8361e-012</td><td>1.7070e-015</td><td>4.3298e-012</td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="4"></td></tr>
<tr><td colspan="2">$f_3(x)=x^2,\ u(x)=\frac{1}{12}\cdot(x-x^4)$</td><td colspan="2">$f_4(x)=x^2+x+1\,,u(x)=-\frac{1}{12}\cdot x\cdot\left(x^3+2x^2+6x-9\right)$</td></tr>
<tr><td>Απόλυτο σφάλμα</td><td>Ποσοστιαίο σφάλμα</td><td>Απόλυτο σφάλμα</td><td>Ποσοστιαίο σφάλμα</td></tr>
<tr><td>9.1593e-016</td><td>4.6190e-012</td><td>6.4115e-015</td><td>3.9841e-012</td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
</table>

Πίνακας 4.2.2.Β. Πειραματικά αποτελέσματα για τις συναρτήσεις f_1 εώς f_4 για την εξίσωση $-u''(x)=f(x)$ με $h = 0.01$.

$f_5(x)=x^3\ ,\ u(x)=\frac{1}{20}\cdot\left(x-x^5\right)$		$f_6(x)=x^3+x^2+x+1\ ,\ u(x)=-\frac{1}{60}\cdot x\cdot\left(3x^4+5x^3+10x^2+30x-48\right)$	
Απόλυτο σφάλμα	**Ποσοστιαίο σφάλμα**	**Απόλυτο σφάλμα**	**Ποσοστιαίο σφάλμα**
5.6899e-016	4.7561e-012	6.9111e-015	3.9860e-012
			

$f_7(x)=e^x\ ,\ u(x)=ex-x-e^x+1$		$f_8(x)=e^x\cdot x\ ,\ u(x)=-e^x(x-2)+2(x-1)-ex$	
Απόλυτο σφάλμα	**Ποσοστιαίο σφάλμα**	**Απόλυτο σφάλμα**	**Ποσοστιαίο σφάλμα**
2.2118e-012	1.0468e-009	1.0058e-011	1.1585e-008
			

Πίνακας 4.2.2.Γ. Πειραματικά αποτελέσματα για τις συναρτήσεις f_5 εώς f_8 για την εξίσωση $-u''(x)=f(x)$ με $h=0.01$.

$f_9(x)=e^x \cdot x^2$, $u(x)=-e^x(x^2-4x+6)+3ex-6x+6$		$f_{10}(x)=e^x \cdot (x^2+x+1)$, $u(x)=-e^x(x^2-3x+5)+3ex-5x+5$	
Απόλυτο σφάλμα	**Ποσοστιαίο σφάλμα**	**Απόλυτο σφάλμα**	**Ποσοστιαίο σφάλμα**
3.7124e-011	7.3845e-008	4.9387e-011	1.3431e-008
			
$f_{11}(x)=e^x \cdot (x^3)$, $u(x)=-e^x(x^3-6x^2+18x-24)+24(x-1)-11ex$		$f_{12}(x)=e^x \cdot (x^3+x^2+x+1)$, $u(x)=-e^x(x^3-5x^2+15x-19)+19(x-1)-8ex$	
Απόλυτο σφάλμα	**Ποσοστιαίο σφάλμα**	**Απόλυτο σφάλμα**	**Ποσοστιαίο σφάλμα**
1.0757e-010	3.0772e-007	1.5687e-010	3.6711e-008
			

Πίνακας 4.2.2.Δ. Πειραματικά αποτελέσματα για τις συναρτήσεις f_9 εώς f_{12} για την εξίσωση $-u''(x)=f(x)$ με $h = 0.01$.

4.3. Εφαρμογή στις εξισώσεις Poisson και Laplace

Για να χρησιμοποιήσουμε την μέθοδο Galerkin στις εξισώσεις Poisson και Laplace κατ 'αρχάς παράγουμε την ασθενή μορφή πολλαπλασιάζοντας με μια δοκιμαστική συνάρτηση $v(x,y)$ και μετέπειτα ολοκληρώνουμε επί του πεδίου ορισμού :

Εξίσωση Laplace

$$-u_{xx}-u_{yy}=0\Rightarrow$$

$$\left(-u_{xx}-u_{xx}\right)v(x,y)=0\Rightarrow$$

$$\iint\left(-u_{xx}-u_{xx}\right)v(x,y)dxdy=0$$

Εξίσωση Poisson

$$-u_{xx}-u_{yy}=f(x,y)\Rightarrow$$

$$\left(-u_{xx}-u_{xx}\right)v(x,y)=f(x,y)\cdot v(x,y)\Rightarrow$$

$$\iint\left(-u_{xx}-u_{xx}\right)v(x,y)dxdy=\iint f(x,y)\cdot v(x,y)dxdy$$

Όπως προαναφέραμε στο τμήμα Α.2. της παραγράφου 2.9 εάν εφαρμόσουμε το θεώρημα του Green στις παραπάνω εξισώσεις προκύπτει ότι :

$$\iint\left(u_x v_x+u_y v_y\right)=0 \quad , \quad \iint\left(u_x v_x+u_y v_y\right)=\iint f(x,y)\cdot v(x,y)dxdy$$

ή ισοδύναμα :

$$\iint\left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial x}+\frac{\partial u}{\partial y}\frac{\partial v}{\partial y}\right)dxdy=0 \quad , \quad \iint\left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial x}+\frac{\partial u}{\partial y}\frac{\partial v}{\partial y}\right)dxdy=\iint f(x,y)\cdot v(x,y)dxdy$$

Κατ΄αναλογία με την αντιμετώπιση στη μια διάσταση, αναζητούμε μια λύση η οποία θα διατυπωθεί ως γραμμικός συνδυασμός των συναρτήσεων βάσης :

$$U=\sum_{j=1}^{N}U_j\varphi_j(x,y)$$

, συνεπώς προκύπτει ότι :

$$\iint\left[\left(\sum_{j=1}^{N}U_j\frac{\partial\varphi_j}{\partial x}\right)\frac{\partial V_i}{\partial x}+\left(\sum_{j=1}^{N}U_j\frac{\partial\varphi_j}{\partial x}\right)\frac{\partial V_i}{\partial y}\right]dxdy=0 \quad ,$$

$$\iint\left[\left(\sum_{j=1}^{N}U_j\frac{\partial\varphi_j}{\partial x}\right)\frac{\partial V_i}{\partial x}+\left(\sum_{j=1}^{N}U_j\frac{\partial\varphi_j}{\partial x}\right)\frac{\partial V_i}{\partial y}\right]dxdy=\iint f(x,y)\cdot V_i(x,y)dxdy$$

Εάν θέσουμε ως:

$$K_{ij}=\iint\left(\frac{\partial\varphi_j}{\partial x}\frac{\partial V_i}{\partial x}+\frac{\partial\varphi_j}{\partial x}\frac{\partial V_i}{\partial x}\right)dxdy, \quad F_i=\iint f(x,y)\cdot V_i(x,y)dxdy$$

, τότε το παραπάνω σύστημα εξισώσεων ανάγεται σε μορφή πινάκων στο σύστημα $KU=F$ το οποίο έχει μοναδική λύση διότι η διγραμμική μορφή και η γραμμική μορφή επληρούν τα κριτήρια του θεωρήματος Lax - Milgram [11] και για τον υπολογισμό των σχετικών ολοκληρωμάτων μπορούμε να χρησιμοποιήσουμε οποιαδήποτε μέθοδο της αριθμητικής ανάλυσης, π.χ. Μέθοδος Simpson.

Εξεταζόμενα προβλήματα

Παρακάτω παραθέτουμε τα προβλήματα που υλοποιήσαμε, όπου για κάθε μια περίπτωση έγιναν υλοποιήσεις για τρείς μορφές στοιχείων, γραμμικά, τετραγωνικά και κυβικά.

	$-u_{xx} - u_{yy} = f(x,y)$			
	$f(x,y)$	Πεδίο ορισμού Ω	Λύση	Συνοριακές συνθήκες
Π1	4	$\{(x,y), x^2+y^2=1\}$	$u(x,y)=1-x^2-y^2$	$u(x,y)=0,\ (x,y)\in\partial\Omega$
Π2	$y^2\sin(xy)+x^2\sin(xy)$	[0,2] x [0,2]	$u(x,y)=\sin(xy)$	$u(0,y)=u(x,0)=0$ $u(1,y)=\sin(2y)$ $u(x,1)=\sin(2x)$

Το σκεπτικό επιλογής των παραπάνω προβλημάτων, είναι ότι είναι ενδεικτικά για την κλάση των προβλημάτων τύπου Laplace και Poission, διότι:

- Στο Π1 έχουμε σταθερή συνάρτηση και μηδενικές συνοριακές συνθήκες, το οποίο είναι υπερσύνολο της κλασικής και πιο εύκολης υπολογιστικά περίπτωσης $f(x,y)=0$ με μηδενικές συνοριακές συνθήκες δηλαδή της εξίσωσης Laplace.
- Στο Π2 έχουμε μη σταθερή συνάρτηση και μη μηδενικές συνοριακές συνθήκες, το οποίο είναι υπερσύνολο της πιο εύκολης υπολογιστικά περίπτωσης με σταθερή συνάρτηση και μη μηδενικές συνοριακές συνθήκες.
- Εξετάζουμε δυο διαφορετικές μορφές πεδίου ορισμού όπου υπάρχουν έντονες διαφορές όπως διαφαίνεται και παρακάτω όταν εφαρμόζουμε πεπερασμένα στοιχεία.

Στο σημείο αυτό να σημειώσουμε ότι ο κώδικας που παρατίθεται στο παράρτημα είναι εύκολα τροποποιήσιμος έτσι ώστε να εφαρμοστεί σε οποιαδήποτε συνάρτηση και για την εφαρμογή της μεθόδου Galerkin στα παραπάνω προβλήματα, βασιστήκαμε :

1. Στην βιβλιοθήκη distmesh που αναπτύχθηκε στο MIT από τον Per-Olof Persson στα πλαίσια της διδακτορικής του διατριβής [21], ο οποίος στην παρούσα φάση θεωρείται από τις πλέον υπολογιστικά συμφέρουσες για παραγωγή πλέγματος και στην εφαρμογή του στην μεθόδο Galerkin, όπως αυτή αναπτύσσεται στην σχετική βιβλιογραφία [5].
2. Ως κύρια προγραμματιστική δομή χρησιμοποιηθηκε η υλοποίηση που αναπτύσσεται για τον παραπάνω αλγόριθμο και οι παραλλαγές του [24] οι οποίες τροποποιήθηκαν και επεκτάθηκαν κατάλληλα έτσι ώστε :
    i. Να υλοποιούνται τετραγωνικά και κυβικά στοιχεία.
    ii. Να αντιμετωπίζονται περιπτώσεις πέραν του τετραγώνου για το πεδίο ορισμού.
    iii. Να αντιμετωπίζονται και μη μηδενικές συνοριακές συνθήκες.
    iv. Επαναληπτική δομή για διευρεύνση της αποτελεσματικότητας της εκλέπτυνσης και παραγωγή γραφήματων.

4.3.1. Πειραματικά αποτελέσματα για τις εξισώσεις Poisson και Laplace

Η βασική δομή του αλγορίθμου που παρατίθεται με πλήρη σχολιασμό στο παράρτημα είναι η εξής:

1. Παράγουμε το πλέγμα χρησιμοποιώντας την βιβλιοθήκη distmesh η οποία μας επιστρέφει τις κορυφές των τριγώγων και τις συντενταγμένες τους, όπου η μέγιστη απόσταση δυο σημείων εντός του τριγώνου ορίζεται από το βήμα h.
2. Κατασκευάζουμε τον τοπικό πίνακα δυσκαμψίας K και το τοπικό διάνυσμα φορτίου F για την εξίσωση Poisson's KU=F για κάθε ένα τριγωνικό στοιχείο λύνοντας το σύστημα και υπολογίζουμε την συνεισφορά κάθε στοιχείου ως το εμβδαδόν του τριγώνου.
3. Κατασκευάζουμε τον καθολικό πίνακα δυσκαμψίας και το καθολικό διάνυσμα φορτίου αθροίζοντας τους αντίστοιχους τοπικούς πίνακες και λύνουμε το σύστημα πινάκων που προκύπτει.
4. Εκτιμούμε τις τιμές της συνάρτησης για τα διάφορα σημεία του πλέγματος και βάσει της προσέγγισης που έχουμε από το προηγούμενο βήμα, βρίσκουμε την νόρμα άπειρο για το σφάλμα.
5. Αποθηκεύουμε σε αρχείο τις τιμές της συνάρτησης, της εκτίμησης και του σφάλματος και αναπαριστούμε γραφικά τα αποτελέσματα.
6. Η παραπάνω δομή του αλγορίθμου εντάχθηκε σε βρόγχο επανάληψης με τις τιμές της εκλέπτυνσης να κυμαίνονται από 0.5 έως και 0.05.

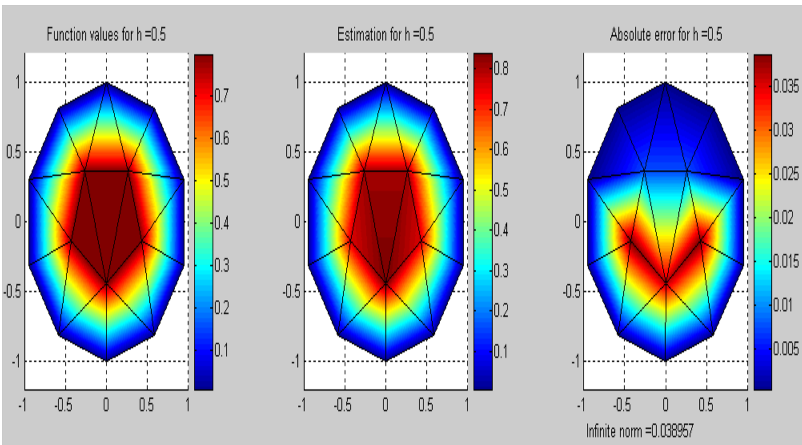
Η εκτέλεση κάθε αρχείου τύπου τύπου m, έχει ως αποτέλεσμα να παραχθεί :

1. Ένα αρχείο τύπου txt στο οποίο αποθηκεύονται τα αριθμητικά αποτελέσματα, με την μορφή που διακρίνεται στο απόσπασμα του αρχείου που παρατίθεται παρακάτω, όπου κατά σειρά καταγράφονται :

- Η τιμή του βήματος
- Η τιμή της συνάρτησης και της εκτίμησης στις κορυφές
- Το απόλυτο σφάλμα για την τιμή στις κορυφές και η νόρμα άπειρο για τα απόλυτα σφάλματα

2. Ένα αρχείο τύπου txt, όπου αποθηκεύονται μόνο οι τιμές για την νόρμα άπειρο.

3. Τρια γραφήματα τα οποία ομαδοποιούνται κατά βήμα εκλέπτυνσης και από αριστερά προς τα δεξιά παρατίθενται η τιμή της συνάρτησης, η τιμή της εκτίμησης και το απόλυτο σφάλμα, με τον χρωματικό κώδικα να κινείται από το κυανό στο ερυθρό καθώς αυξάνει η αριθμητική τιμή.




```
--------------------------------------------------------------------------------
h =  0.50
--------------------------------------------------------------------------------
Πραγματική   Εκτίμηση     Απόλυτο Σφάλμα
0.0000000014 0.0000000000 0.0000000014
0.0000000000 0.0000000000 0.0000000000
-0.0000000015 0.0000000000 0.0000000015
-0.0000000008 0.0000000000 0.0000000008
0.7995701313 0.7606127961 0.0389573352
0.8031346698 0.8105542629 0.0074195931
0.0000000000 0.0000000000 0.0000000000
0.0000000000 0.0000000000 0.0000000000
0.8061408123 0.8426423337 0.0365015214
0.8031346686 0.8105542621 0.0074195935
0.7995701314 0.7606127964 0.0389573350
-0.0000000008 0.0000000000 0.0000000008
-0.0000000015 0.0000000000 0.0000000015
0.0000000000 0.0000000000 0.0000000000
-0.0000000003 0.0000000000 0.0000000003

Νόρμα άπειρο απολύτων σφαλμάτων =  0.0389573352
--------------------------------------------------------------------------------
```

Π1	Νόρμα άπειρο		
h	**Γραμμικά στοιχεία**	**Τετραγωνικά στοιχεία**	**Κυβικά στοιχεία**
0.5	3,90E-02	9,59E-02	8,52E-02
0.4	6,39E-03	6,83E-02	6,07E-02
0.3	1,50E-03	3,60E-02	3,20E-02
0.2	7,34E-03	1,16E-02	1,03E-02
0.1	1,61E-03	3,84E-03	3,41E-03
0.09	1,31E-03	2,60E-03	2,31E-03
0.08	7,86E-04	1,94E-03	1,72E-03
0.07	6,48E-04	1,57E-03	1,40E-03
0.06	5,23E-04	1,20E-03	1,07E-03
0.05	3,85E-04	8,32E-04	7,40E-04

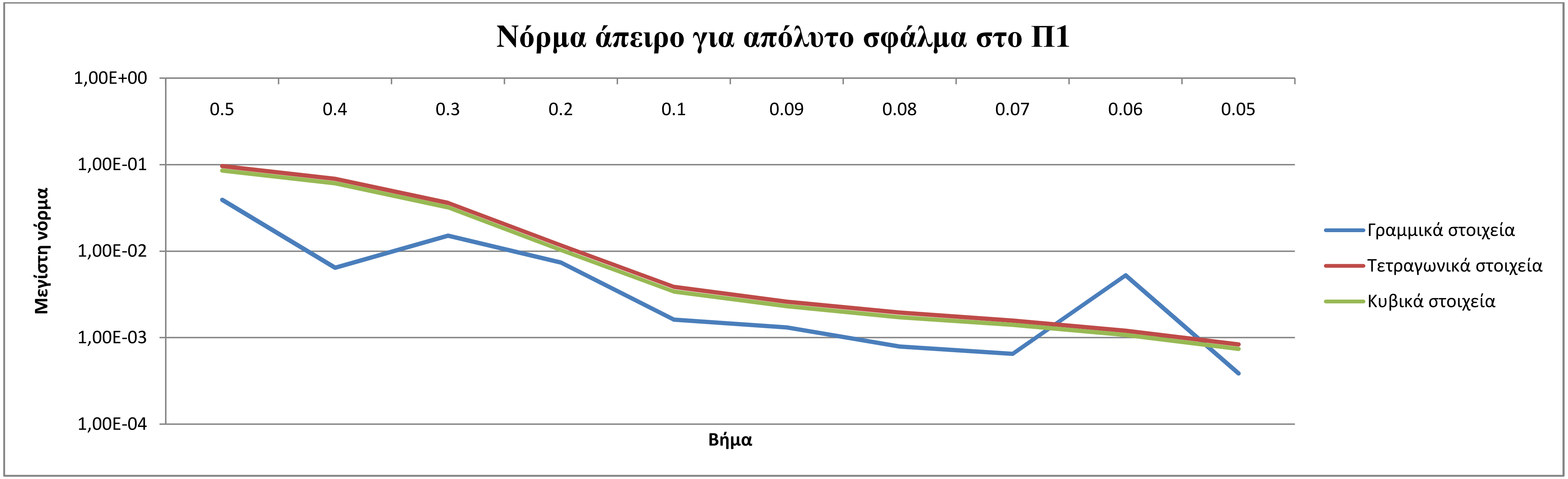
Πίνακας 4.3.1.Α. Τιμές της νόρμα άπειρο στο Π1, για διαφορετικές τιμές εκλέπτυνσης και τύπους στοιχείων.

Π2	Νόρμα άπειρο		
h	**Γραμμικά στοιχεία**	**Τετραγωνικά στοιχεία**	**Κυβικά στοιχεία**
0.5	2,50E-02	7,42E-03	1,38E-03
0.4	2,44E-02	1,52E-03	2,90E-04
0.3	1,70E-02	1,78E-03	1,26E-04
0.2	4,88E-03	2,59E-04	3,33E-05
0.1	1,11E-03	4,35E-05	1,75E-06
0.09	9,04E-04	3,42E-05	1,23E-06
0.08	5,45E-04	1,87E-05	6,96E-07
0.07	5,58E-04	1,31E-05	4,70E-07
0.06	3,80E-04	1,43E-05	2,29E-07
0.05	2,71E-04	6,09E-06	1,05E-07

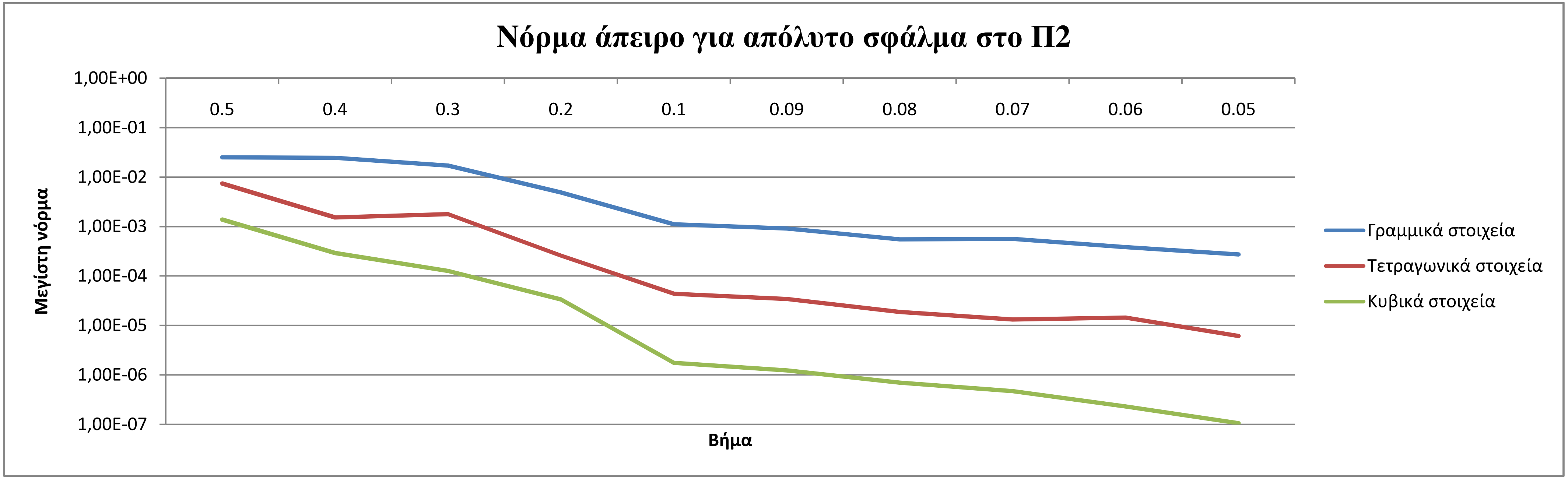
Πίνακας 4.3.1.Β. Τιμές της νόρμα άπειρο στο Π2, για διαφορετικές τιμές εκλέπτυνσης και τύπους στοιχείων.

Στις παρακάτω σελίδες κατά σειρά παρατίθενται:

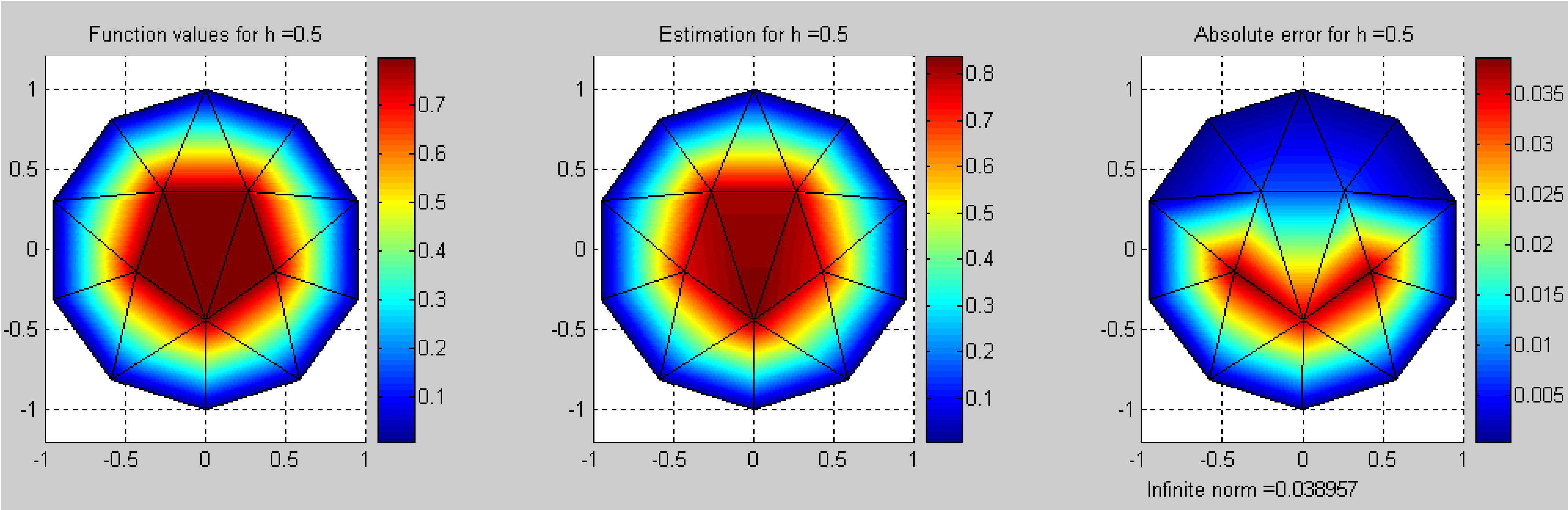
- Συγκριτικό γράφημα για τους τρείς τύπους στοιχείων στο πρόβλημα Π1, σε λογαριθμική κλίμακα για την νόρμα άπειρο για διαφορετικές τιμές της εκλέπτυνσης.
- Συγκριτικό γράφημα για τους τρείς τύπους στοιχείων στο πρόβλημα Π2, σε λογαριθμική κλίμακα για την νόρμα άπειρο για διαφορετικές τιμές της εκλέπτυνσης.
- 180 γραφήματα σε τριάδες της μορφής που προαναφέραμε, με την εξής σειρά :
    - Γραμμικά στοιχεία για το Π1, για τιμές εκλέπτυνσης από 0.5 έως και 0.05 .
    - Τετραγωνικά στοιχεία για το Π1, για τιμές εκλέπτυνσης από 0.5 έως και 0.05 .
    - Κυβικά στοιχεία για το Π1, για τιμές εκλέπτυνσης από 0.5 έως και 0.05 .
    - Γραμμικά στοιχεία για το Π2, για τιμές εκλέπτυνσης από 0.5 έως και 0.05 .
    - Τετραγωνικά στοιχεία για το Π2, για τιμές εκλέπτυνσης από 0.5 έως και 0.05 .
    - Κυβικά στοιχεία για το Π2, για τιμές εκλέπτυνσης από 0.5 έως και 0.05 .



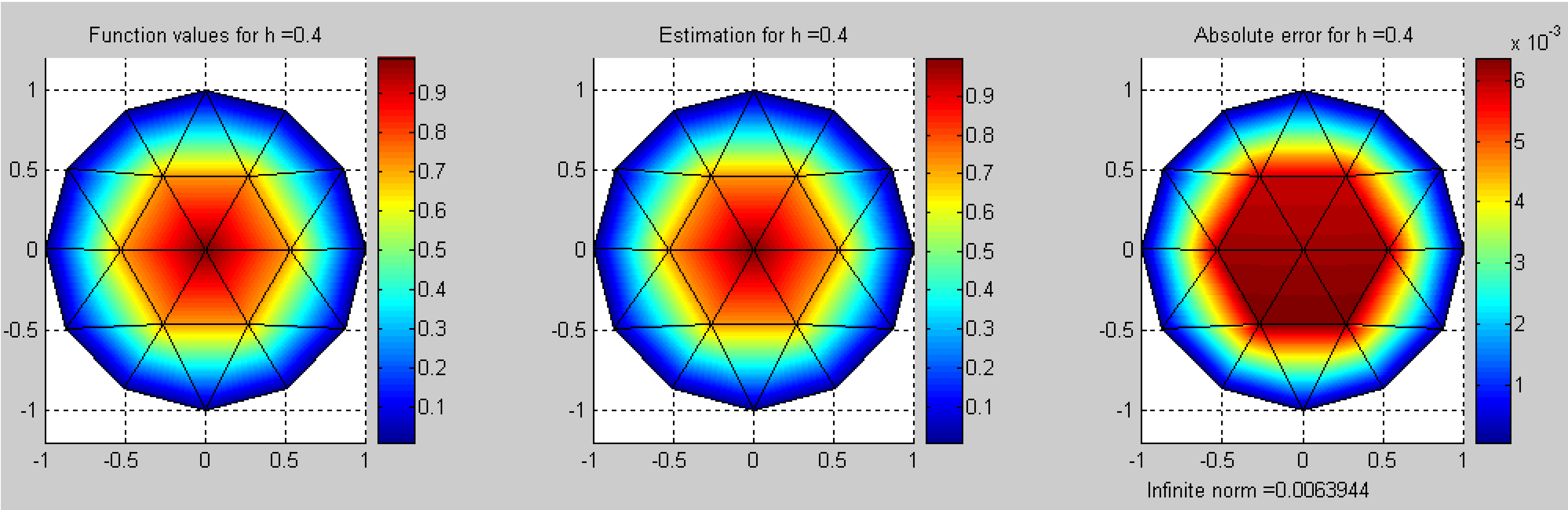
Τιμές της μεγίστης νόρμας σε λογαριθμική κλίμακα για το απόλυτο σφάλμα, με τρείς τύπους πεπερασμένων στοιχείων και διάφορες τιμές βήματος στο Π1.



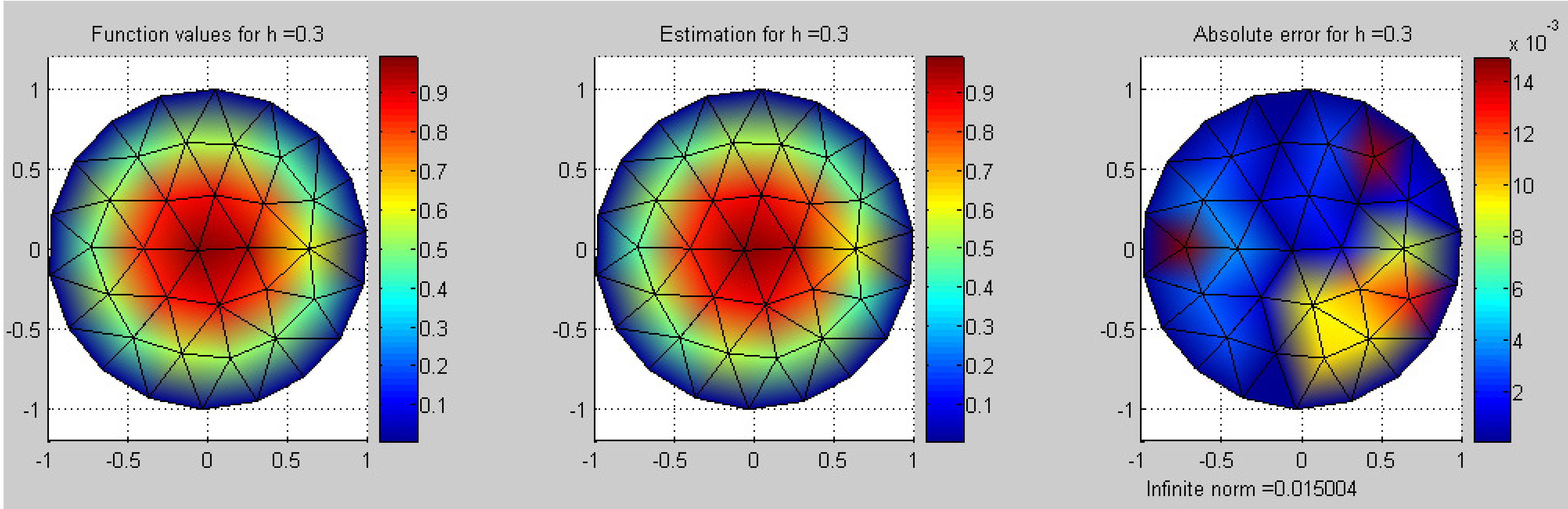
Τιμές της μεγίστης νόρμας σε λογαριθμική κλίμακα για το απόλυτο σφάλμα, με τρείς τύπους πεπερασμένων στοιχείων και διάφορες τιμές βήματος στο Π2.



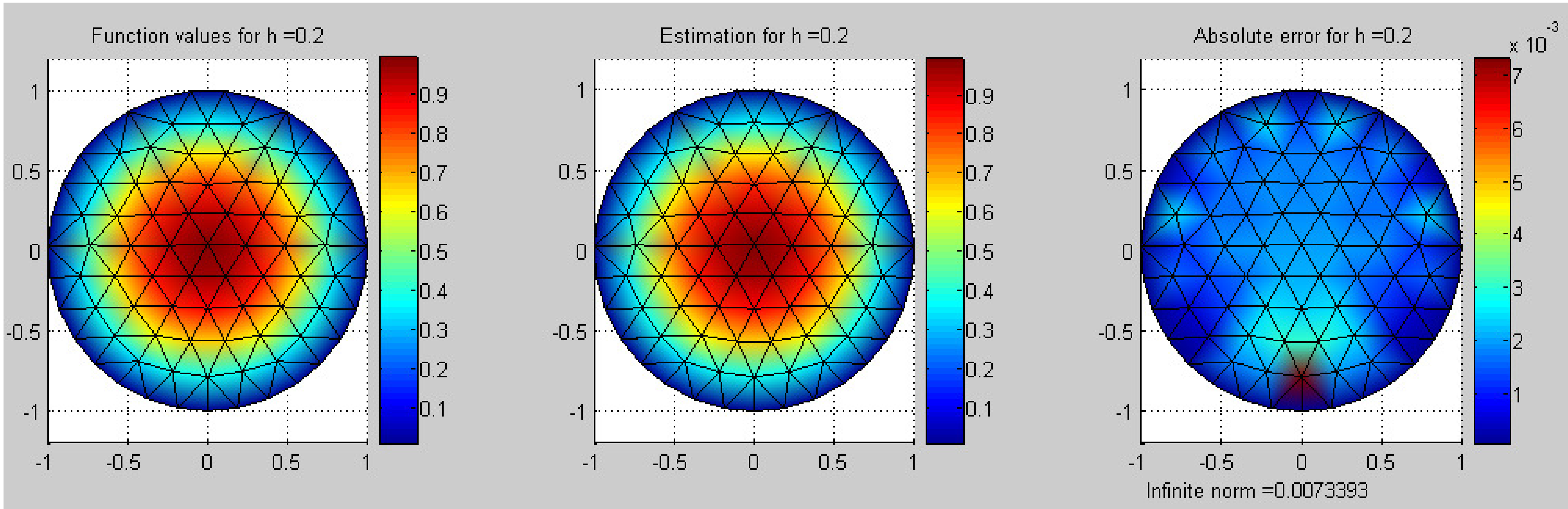
Πειραματικά αποτελέσματα με γραμμικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.5 στο Π1.



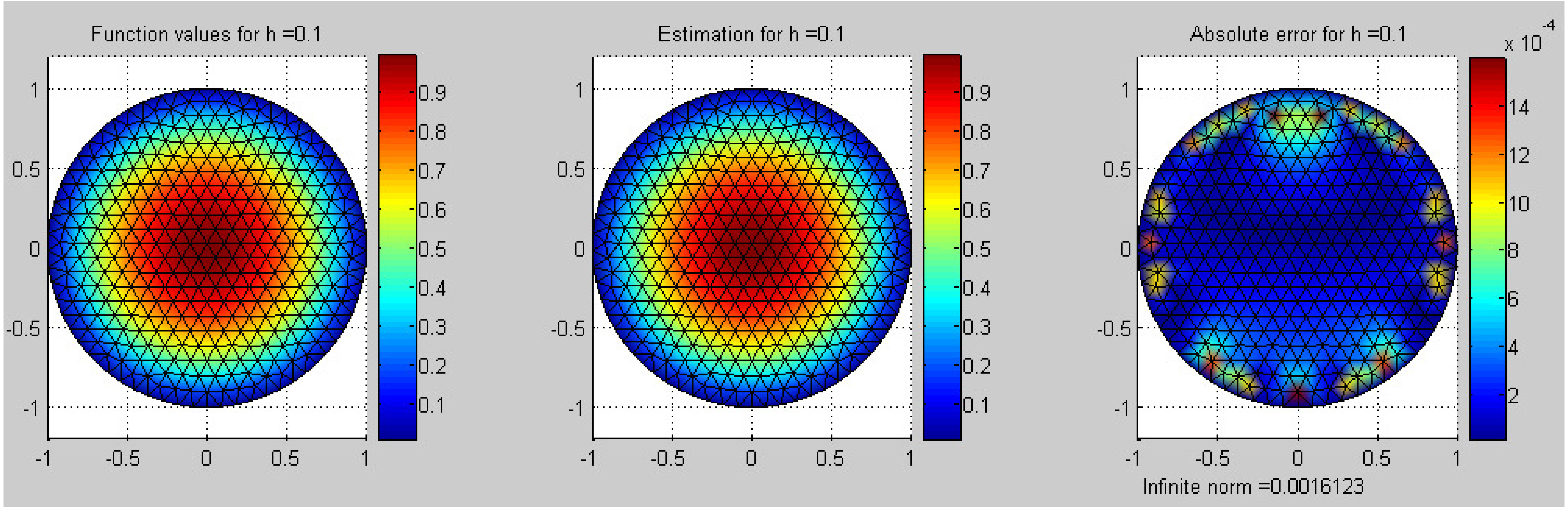
Πειραματικά αποτελέσματα με γραμμικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.4 στο Π1.



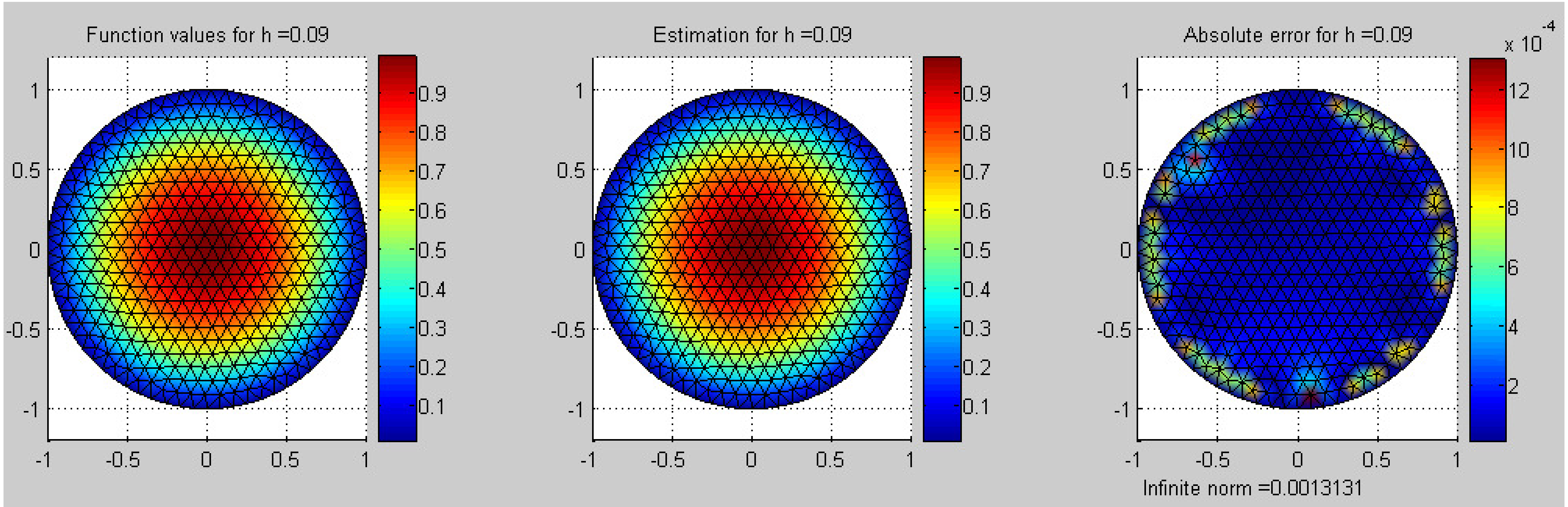
Πειραματικά αποτελέσματα με γραμμικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.3 στο Π1.



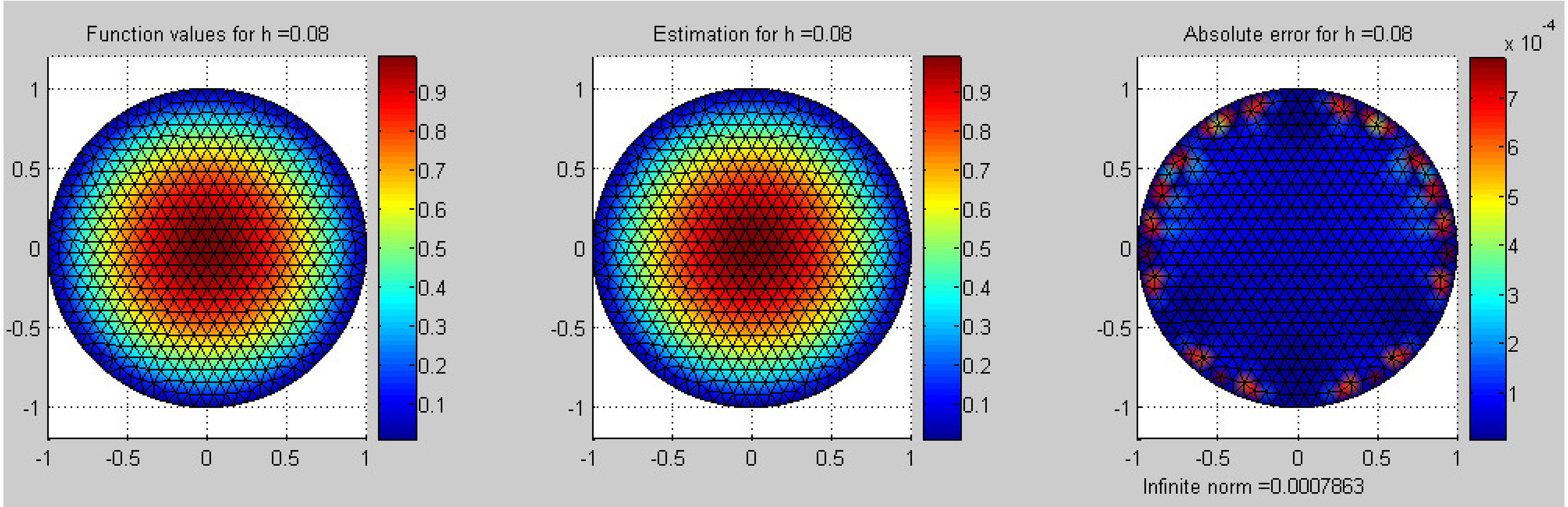
Πειραματικά αποτελέσματα με γραμμικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.2 στο Π1.



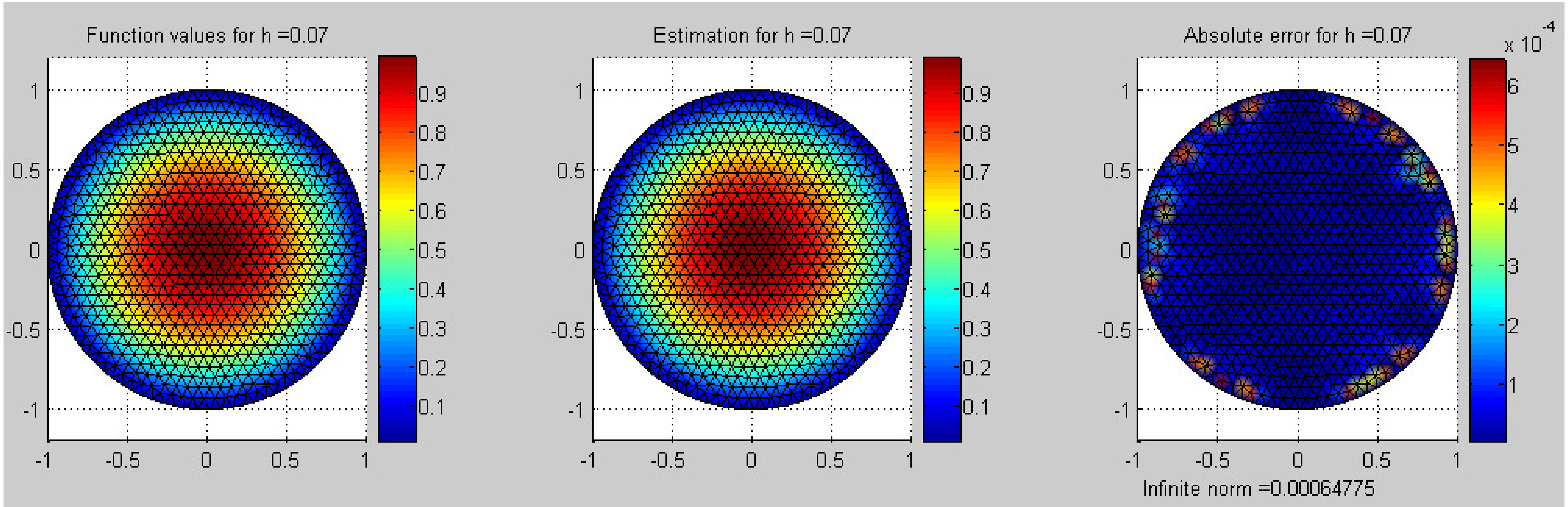
Πειραματικά αποτελέσματα με γραμμικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.1 στο Π1.



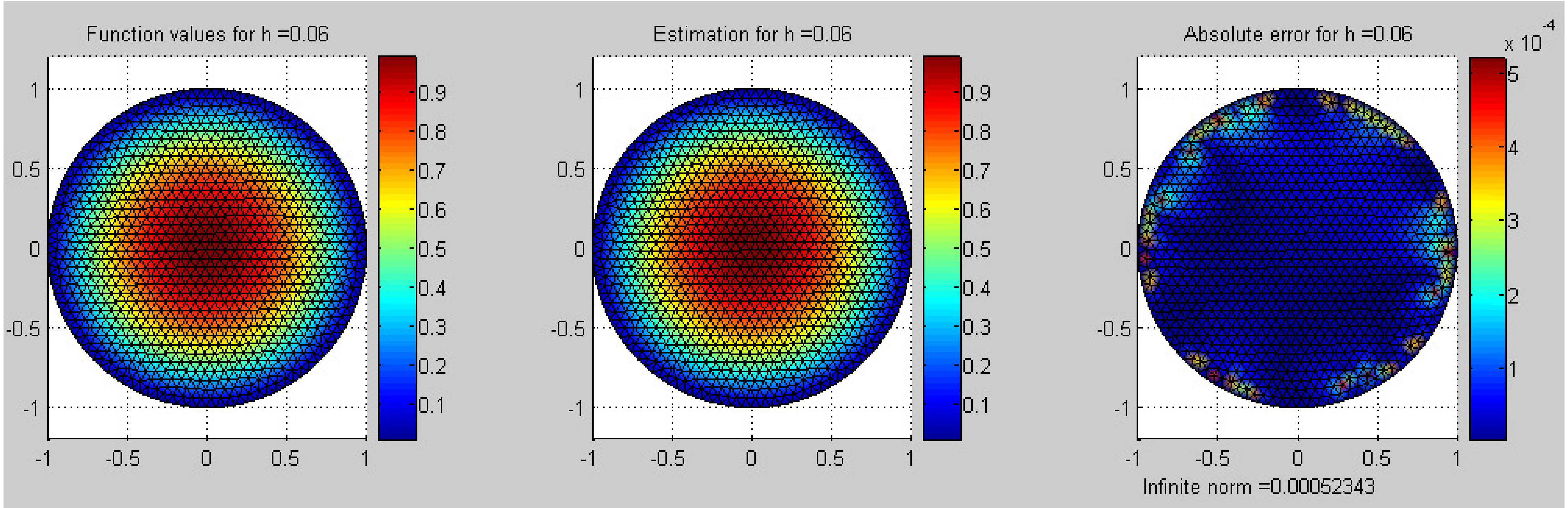
Πειραματικά αποτελέσματα με γραμμικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.09 στο Π1.



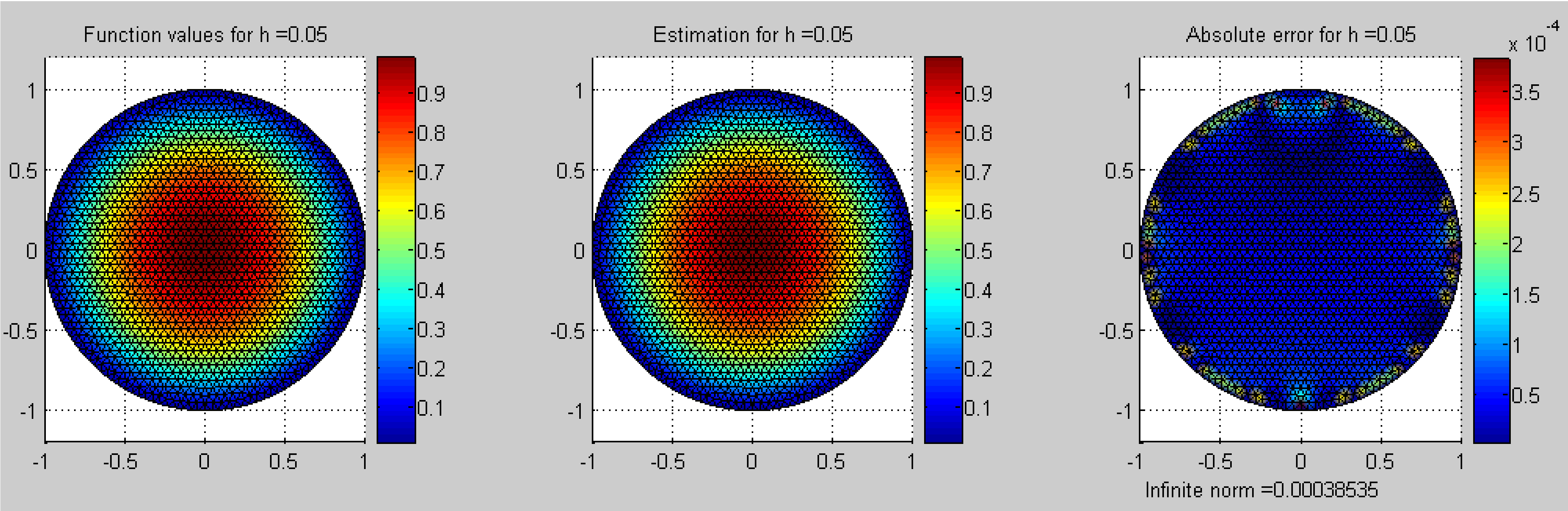
Πειραματικά αποτελέσματα με γραμμικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.08 στο Π1.



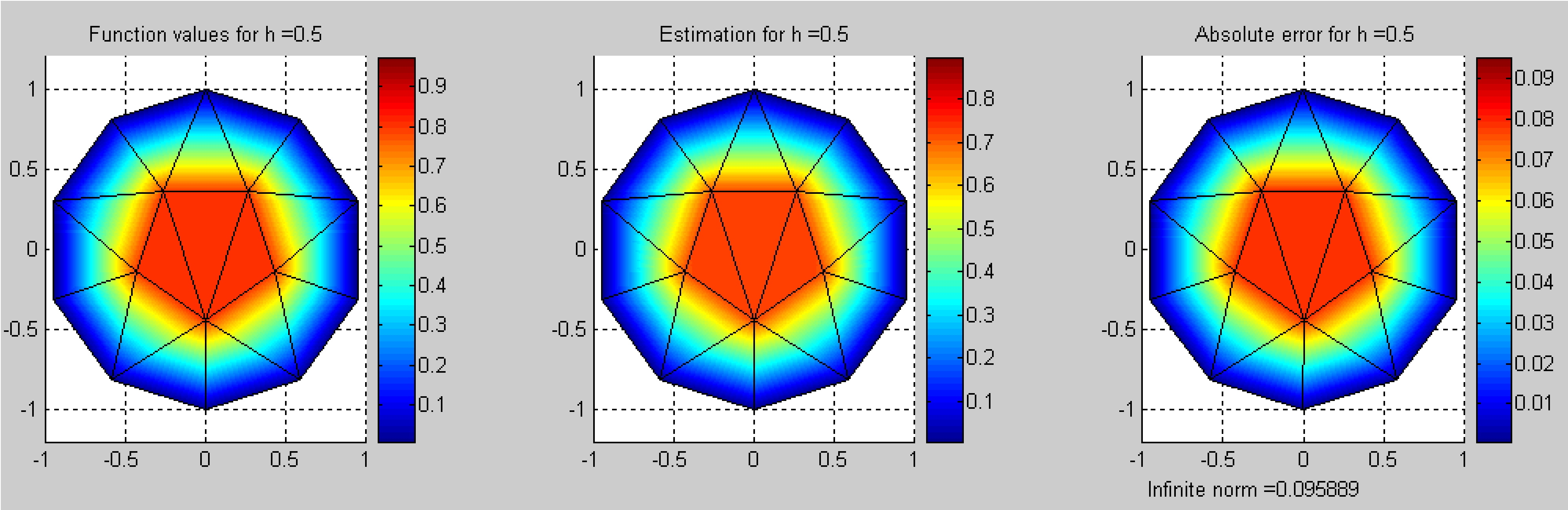
Πειραματικά αποτελέσματα με γραμμικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.07 στο Π1.



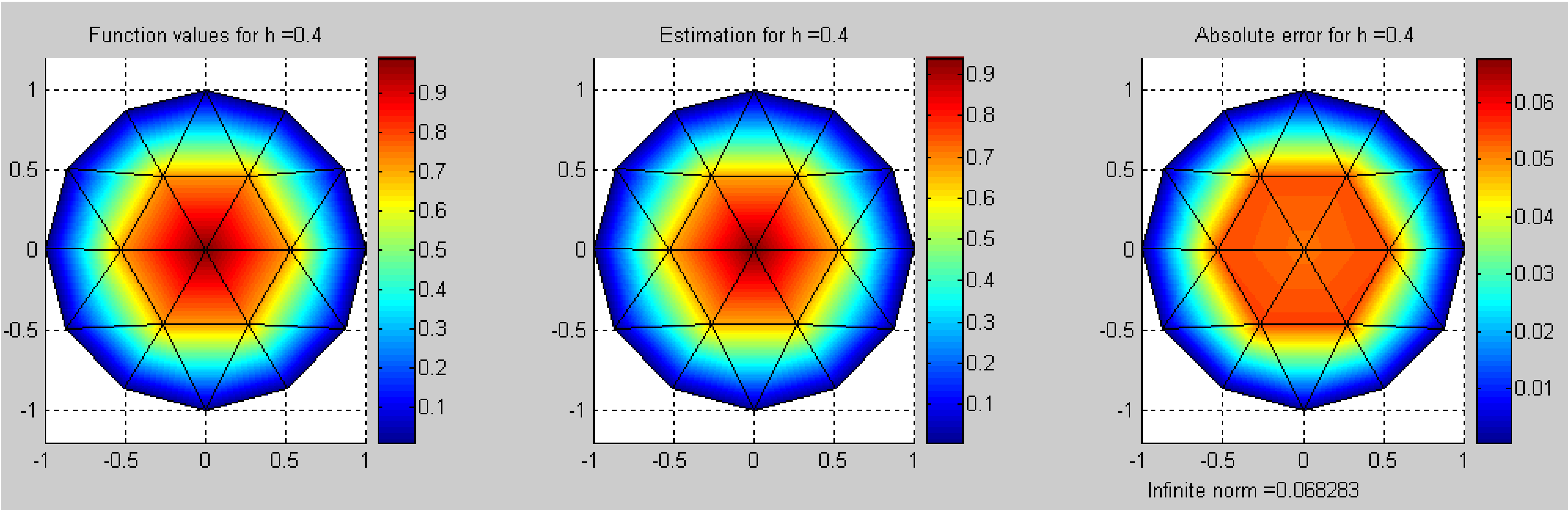
Πειραματικά αποτελέσματα με γραμμικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.06 στο Π1.



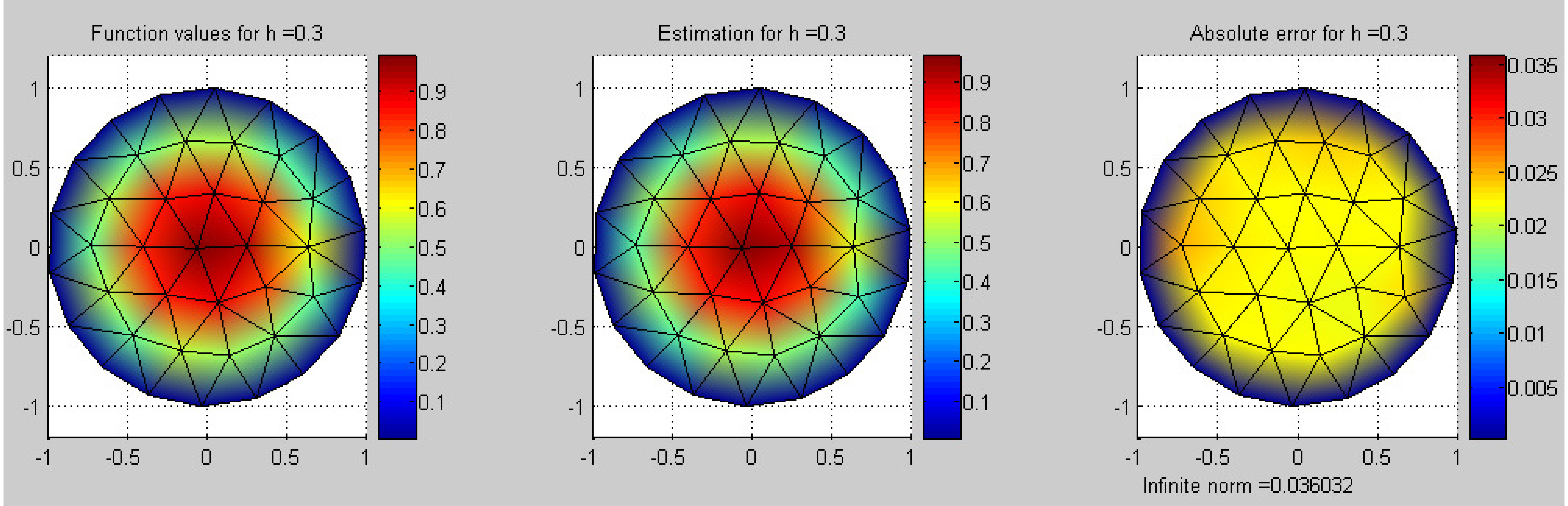
Πειραματικά αποτελέσματα με γραμμικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.05 στο Π1.



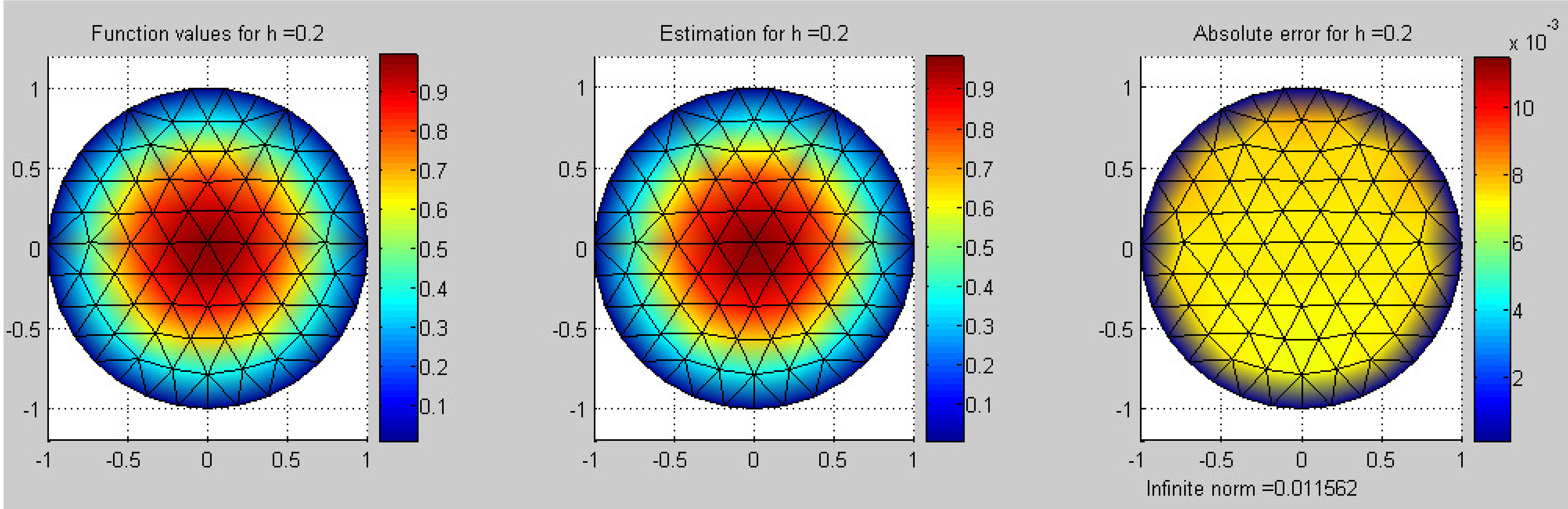
Πειραματικά αποτελέσματα με τετραγωνικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.5 στο Π1.



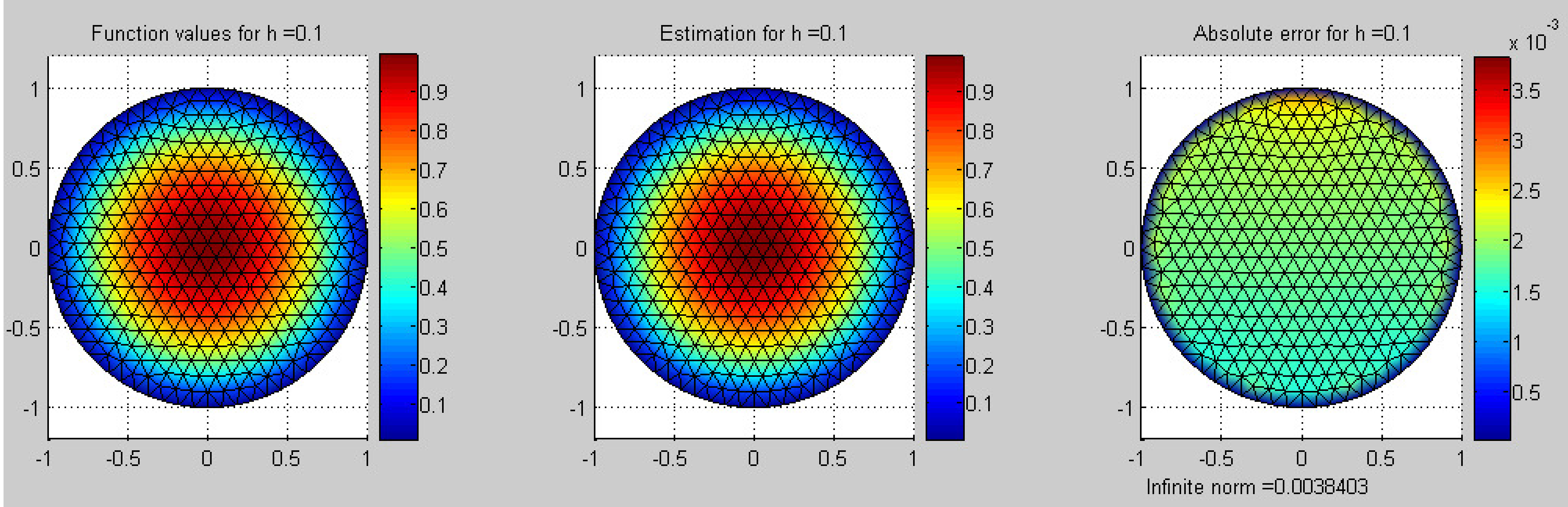
Πειραματικά αποτελέσματα με τετραγωνικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.4 στο Π1.



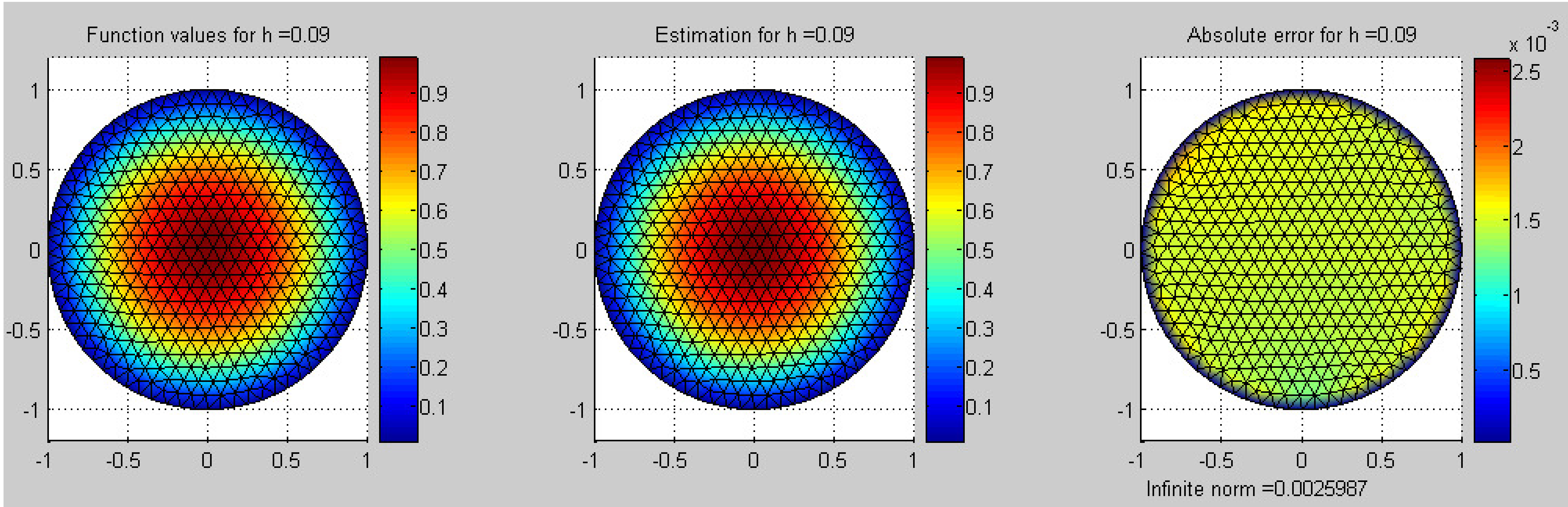
Πειραματικά αποτελέσματα με τετραγωνικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.3 στο Π1.



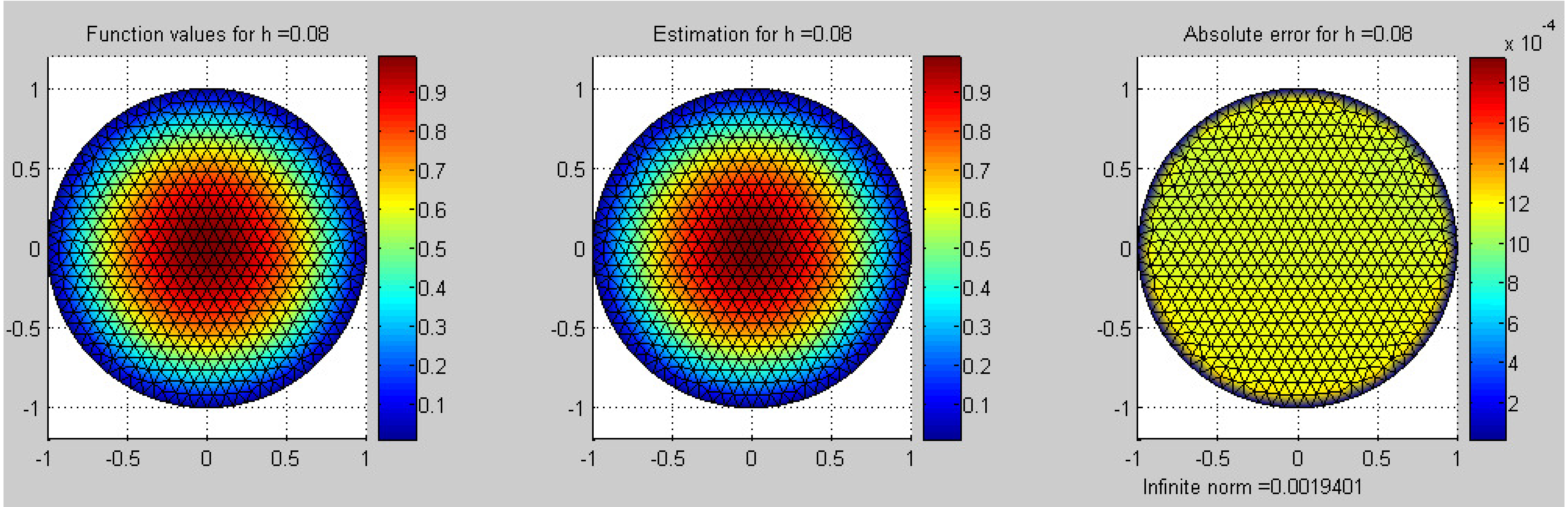
Πειραματικά αποτελέσματα με τετραγωνικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.2 στο Π1.



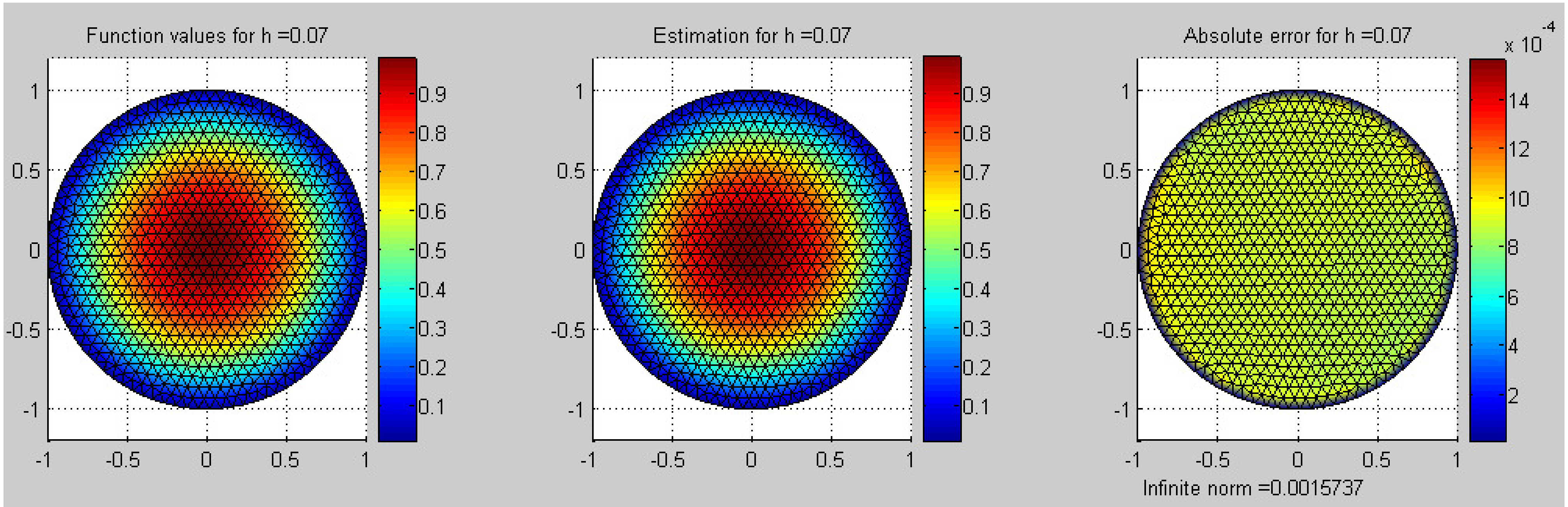
Πειραματικά αποτελέσματα με τετραγωνικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.1 στο Π1.



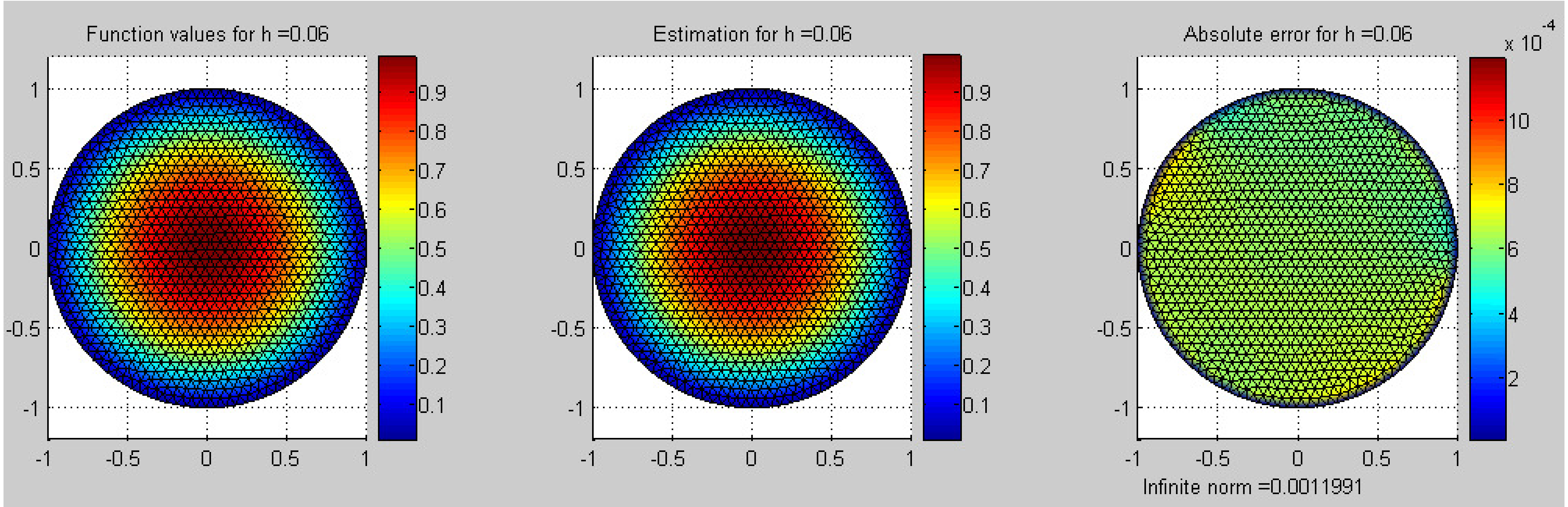
Πειραματικά αποτελέσματα με τετραγωνικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.09 στο Π1.



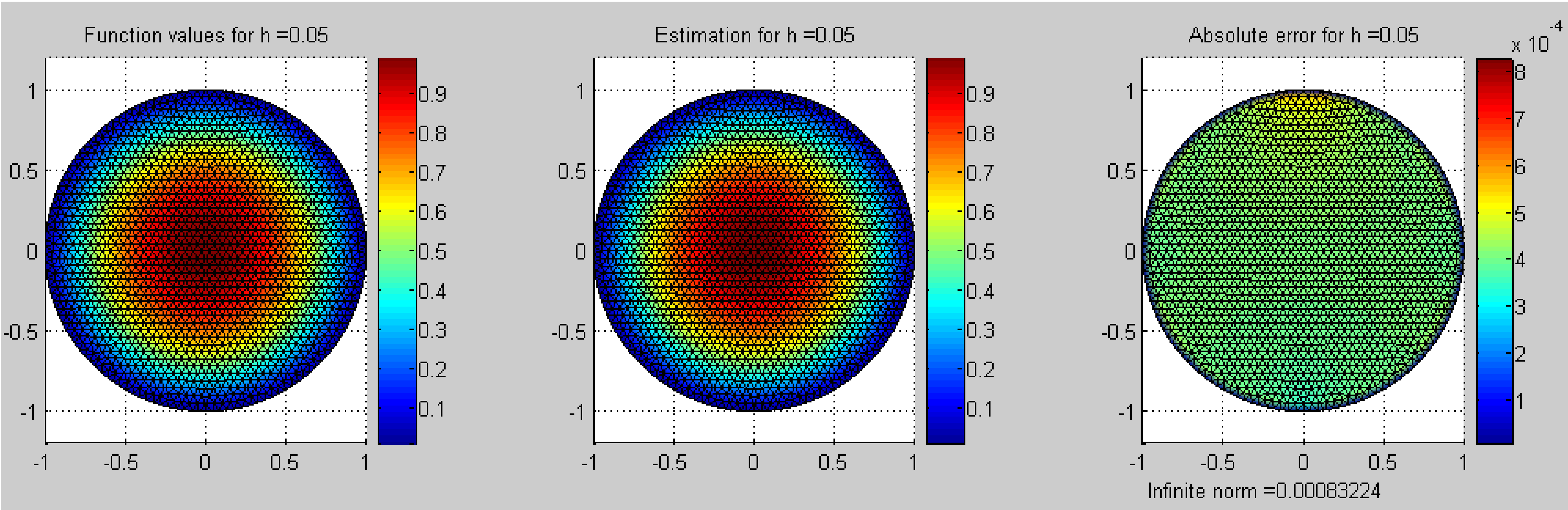
Πειραματικά αποτελέσματα με τετραγωνικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.08 στο Π1.



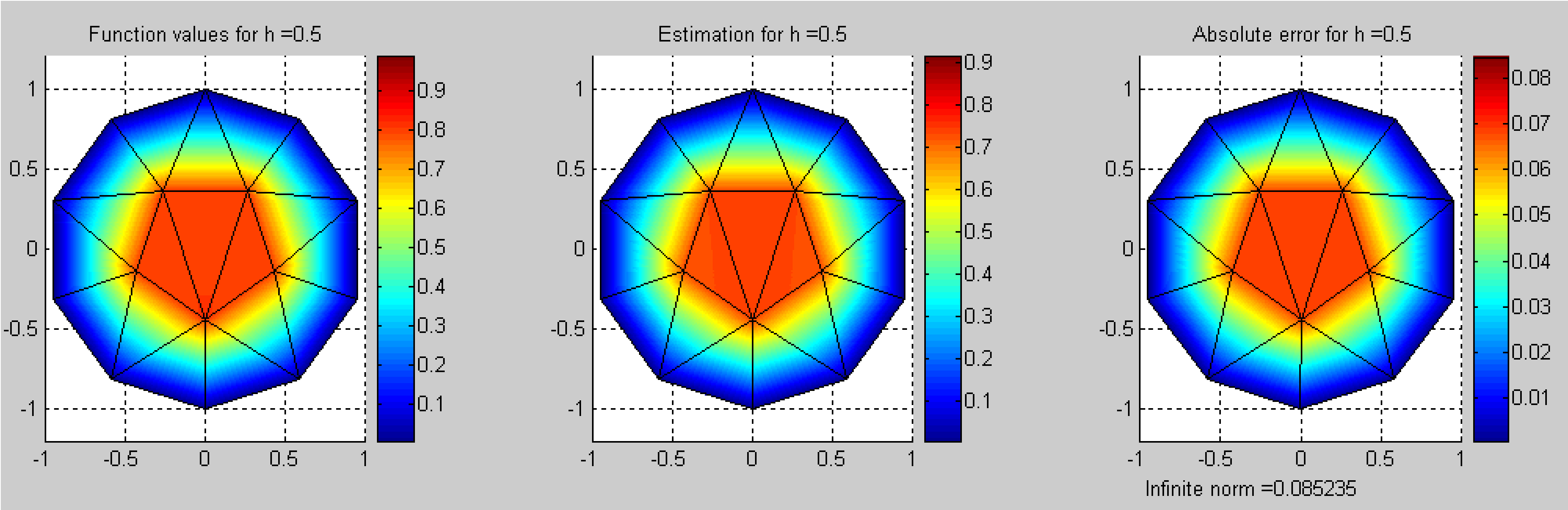
Πειραματικά αποτελέσματα με τετραγωνικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.07 στο Π1.



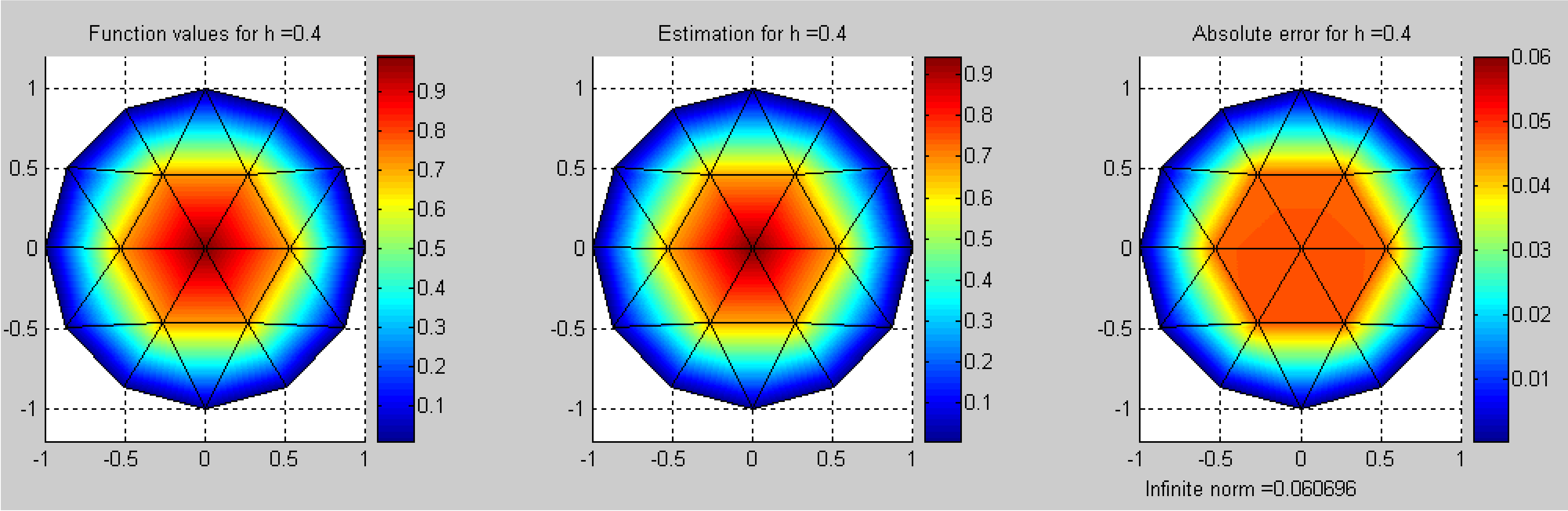
Πειραματικά αποτελέσματα με τετραγωνικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.06 στο Π1.



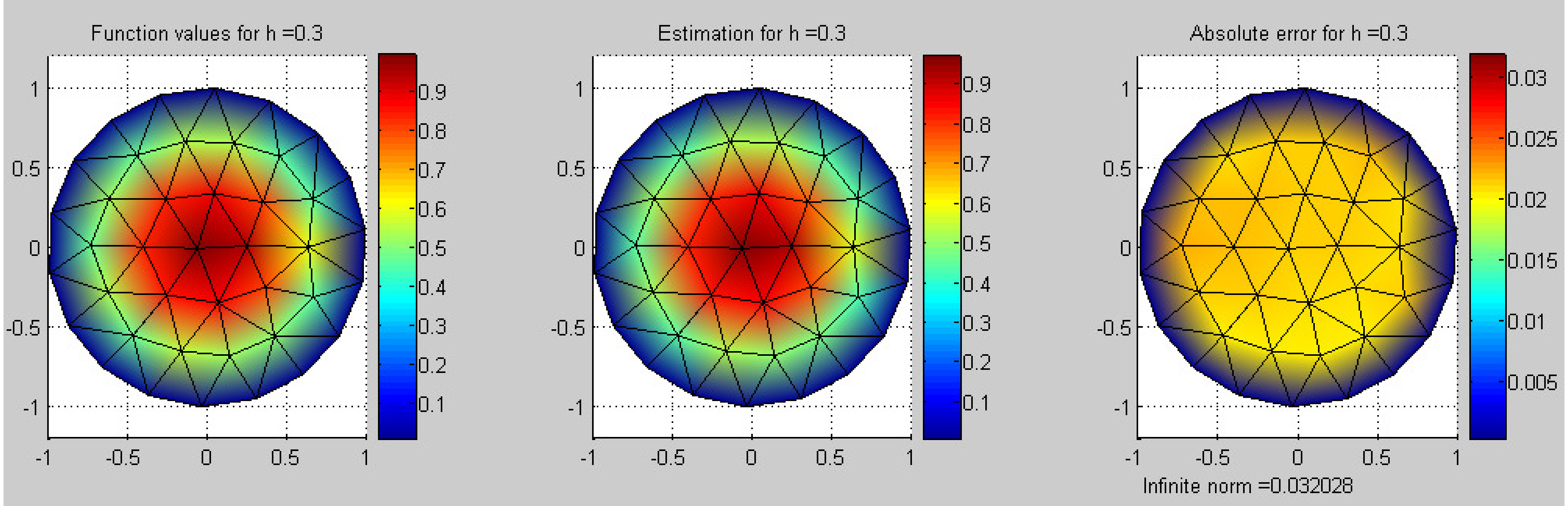
Πειραματικά αποτελέσματα με τετραγωνικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.05 στο Π1.



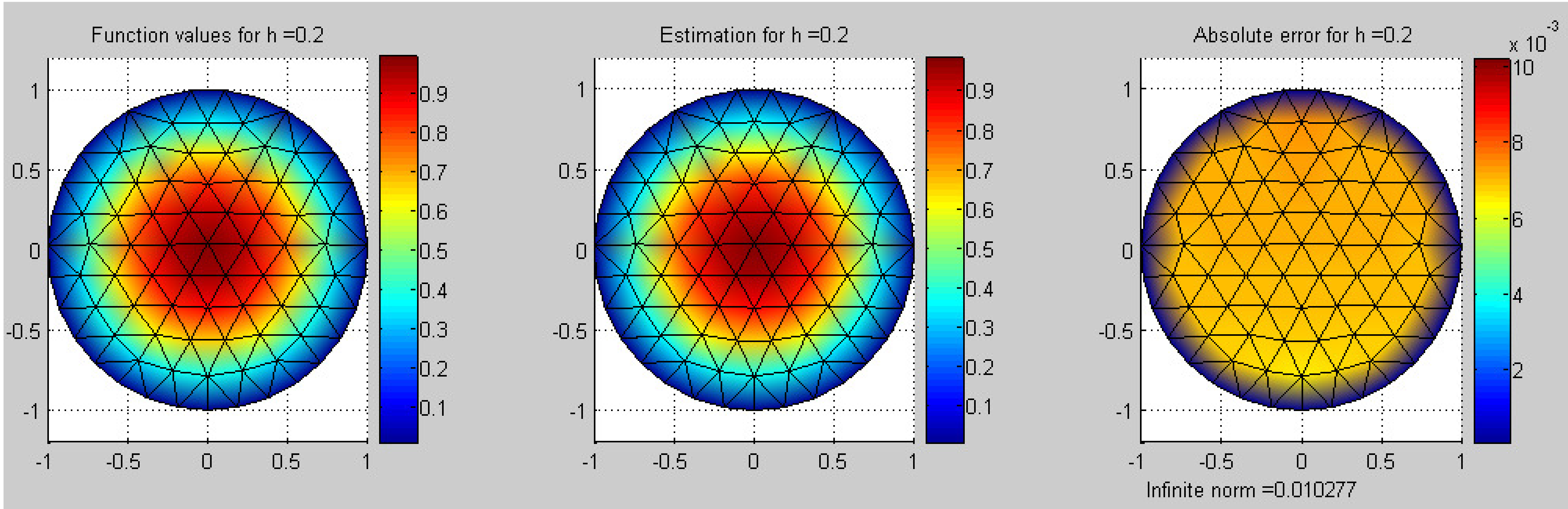
Πειραματικά αποτελέσματα με κυβικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.5 στο Π1.



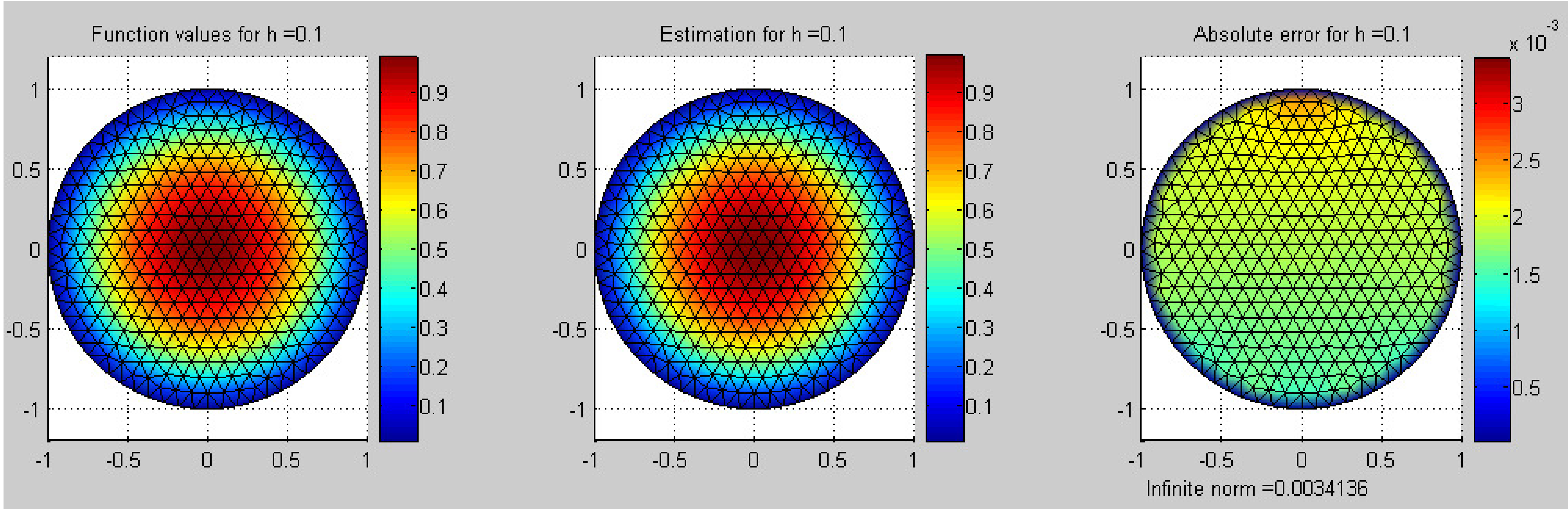
Πειραματικά αποτελέσματα με κυβικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.4 στο Π1.



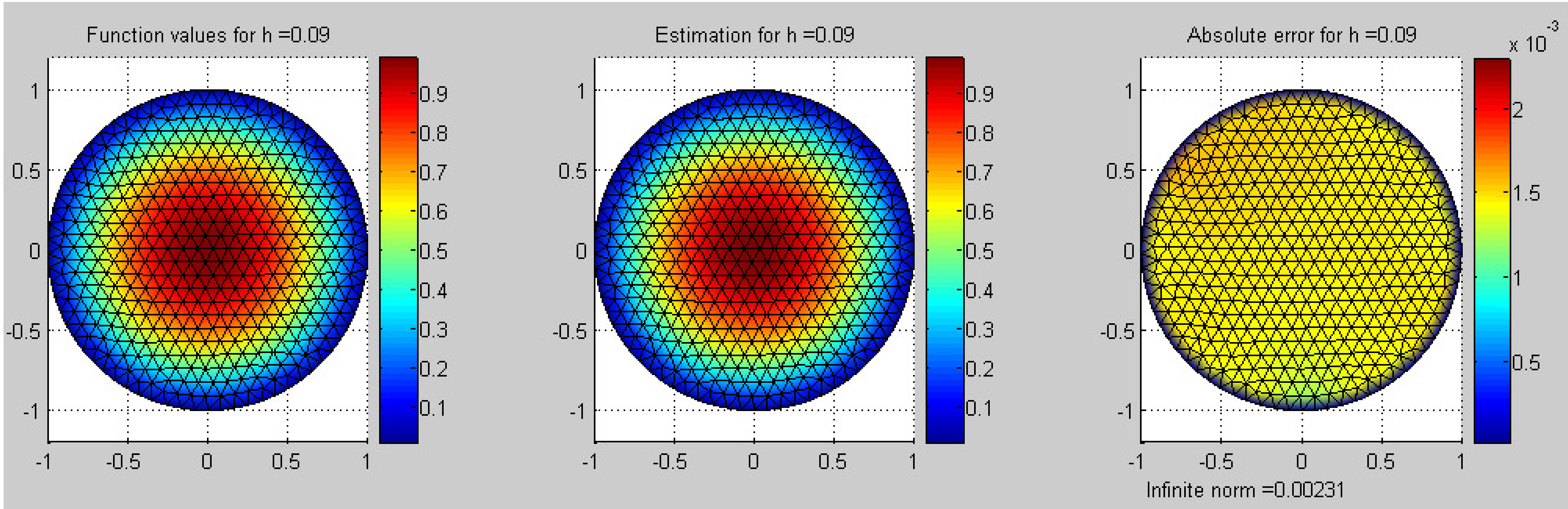
Πειραματικά αποτελέσματα με κυβικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.3 στο Π1.



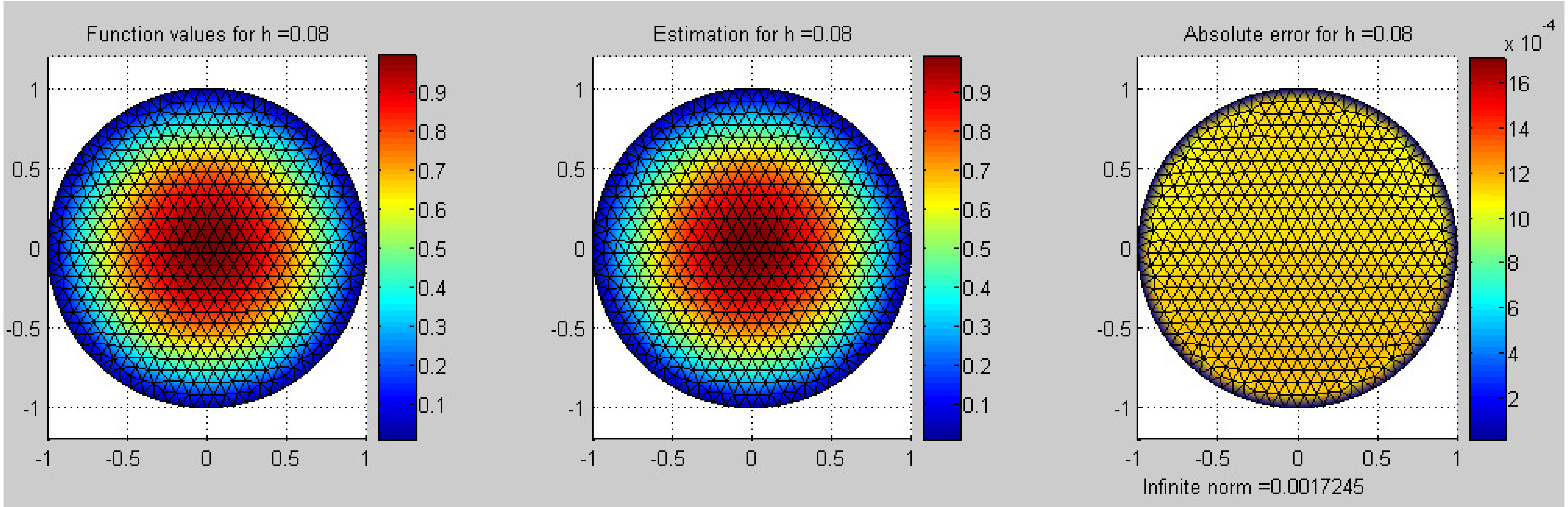
Πειραματικά αποτελέσματα με κυβικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.2 στο Π1.



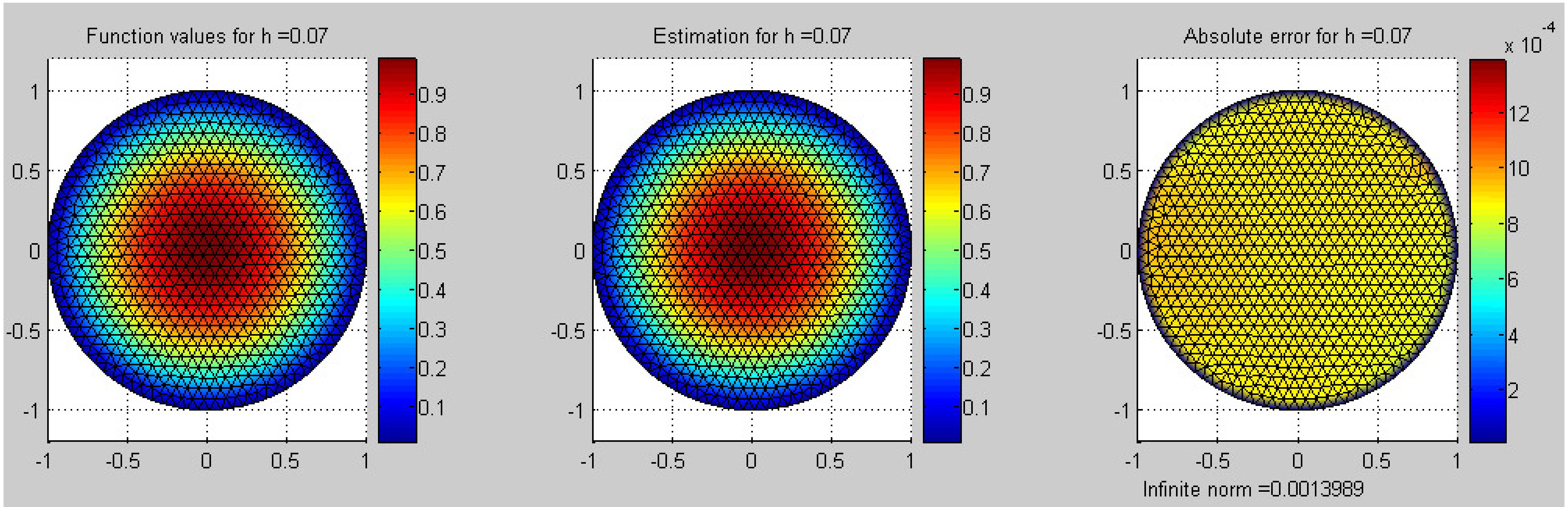
Πειραματικά αποτελέσματα με κυβικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.1 στο Π1.



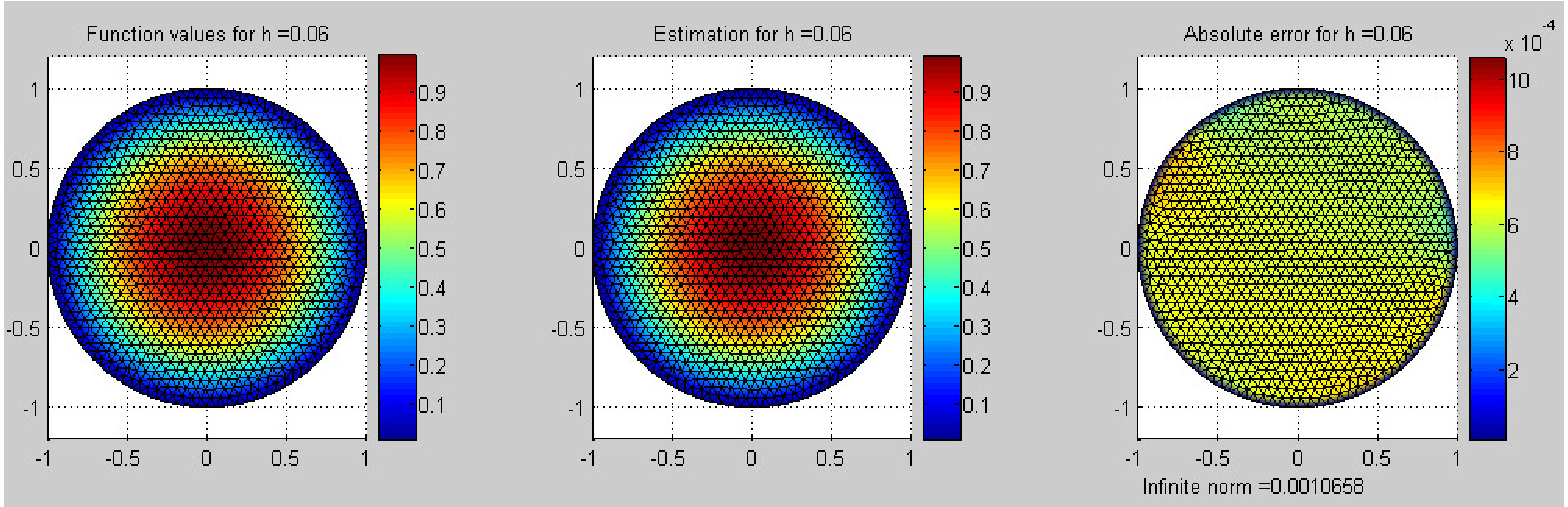
Πειραματικά αποτελέσματα με κυβικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.09 στο Π1.



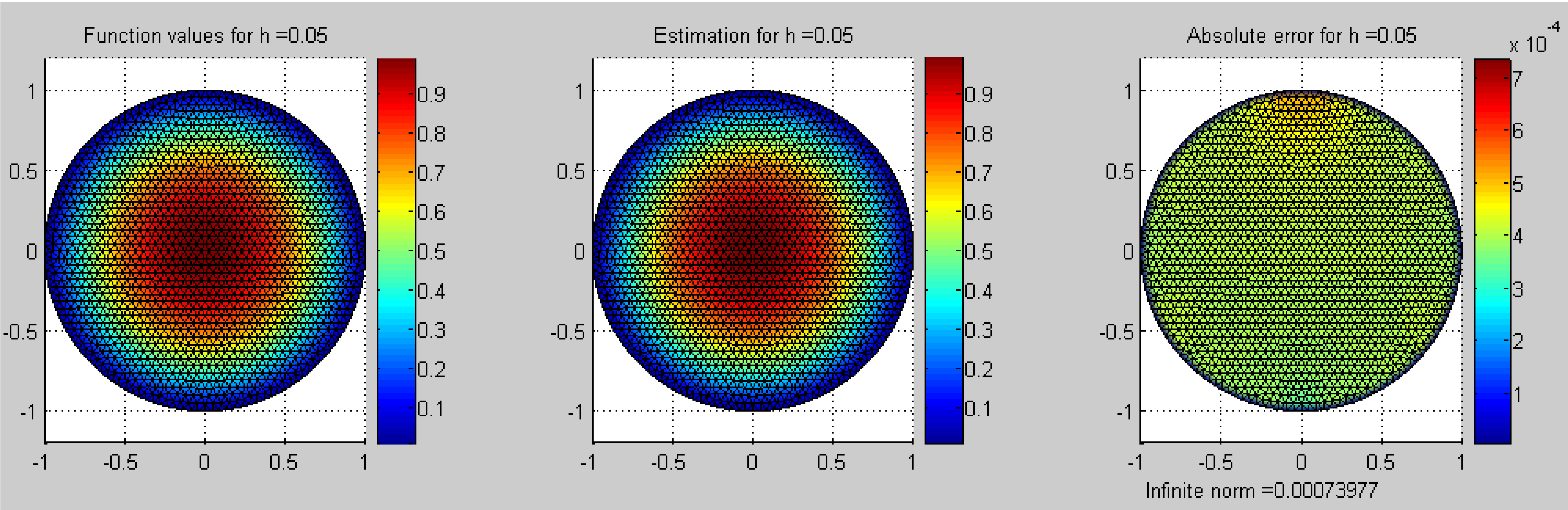
Πειραματικά αποτελέσματα με κυβικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.08 στο Π1.



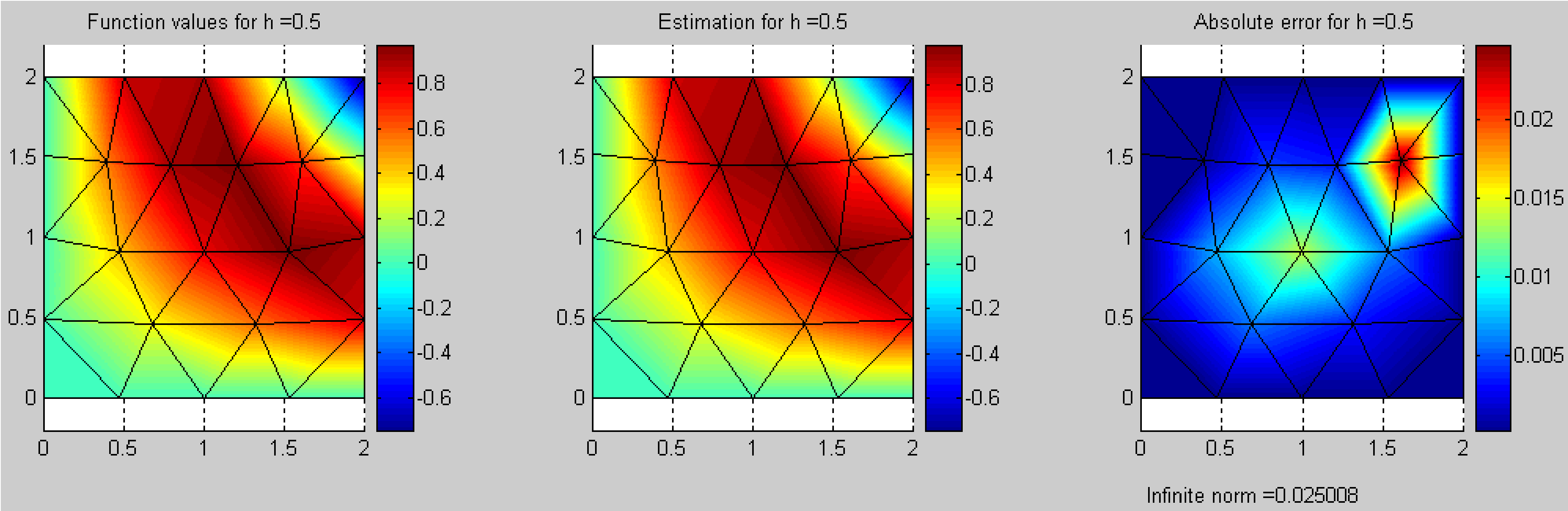
Πειραματικά αποτελέσματα με κυβικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.07 στο Π1.



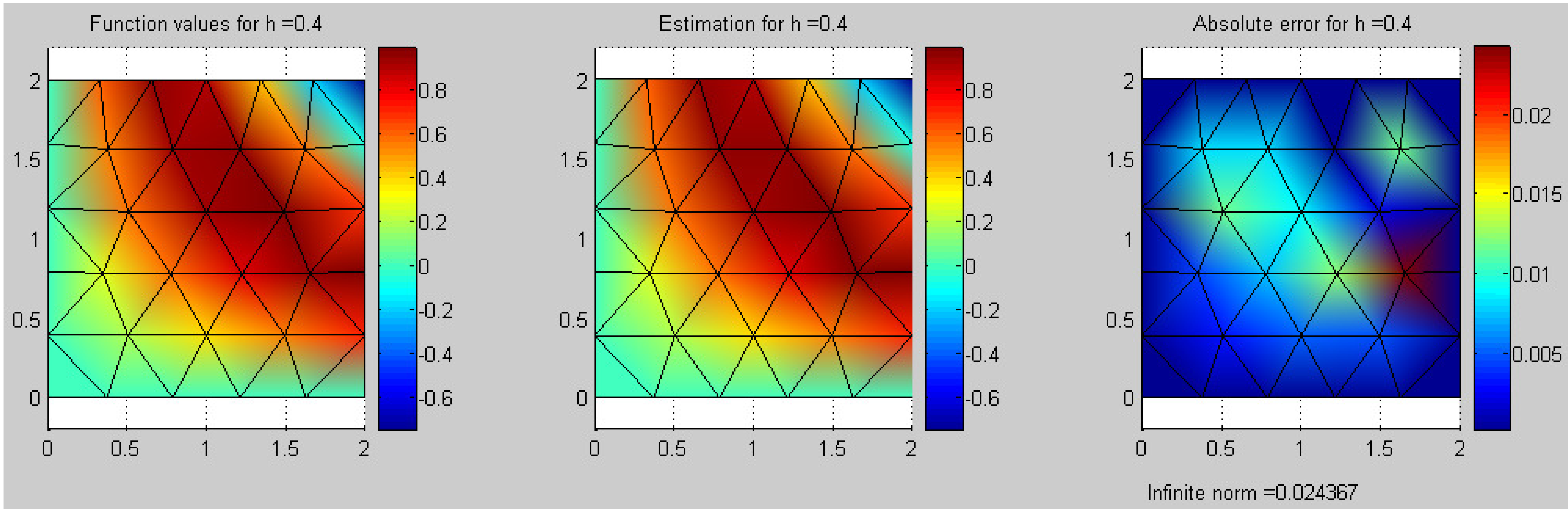
Πειραματικά αποτελέσματα με κυβικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.06 στο Π1.



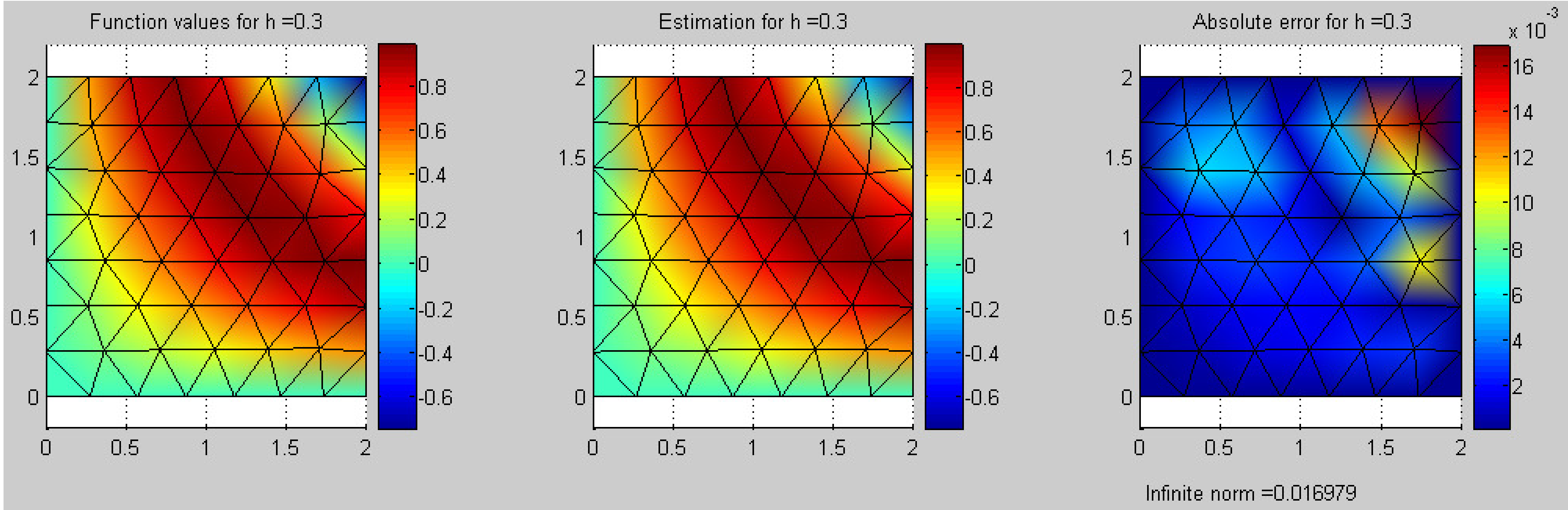
Πειραματικά αποτελέσματα με κυβικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.05 στο Π1.



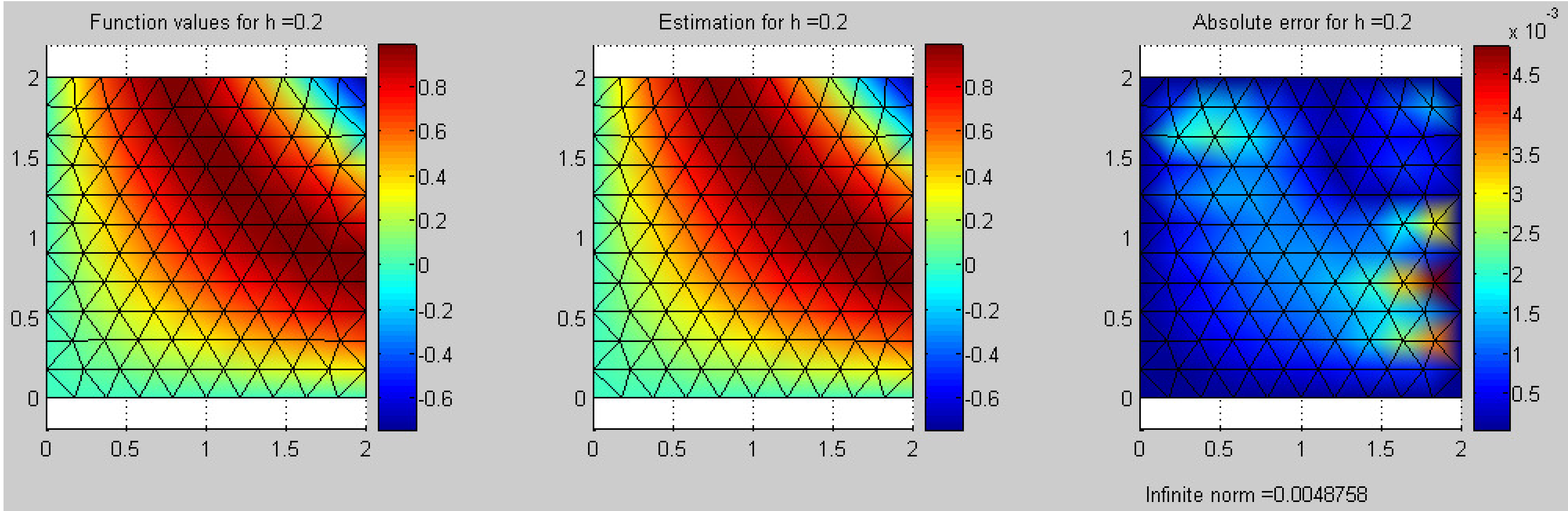
Πειραματικά αποτελέσματα με γραμμικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.5 στο Π2.



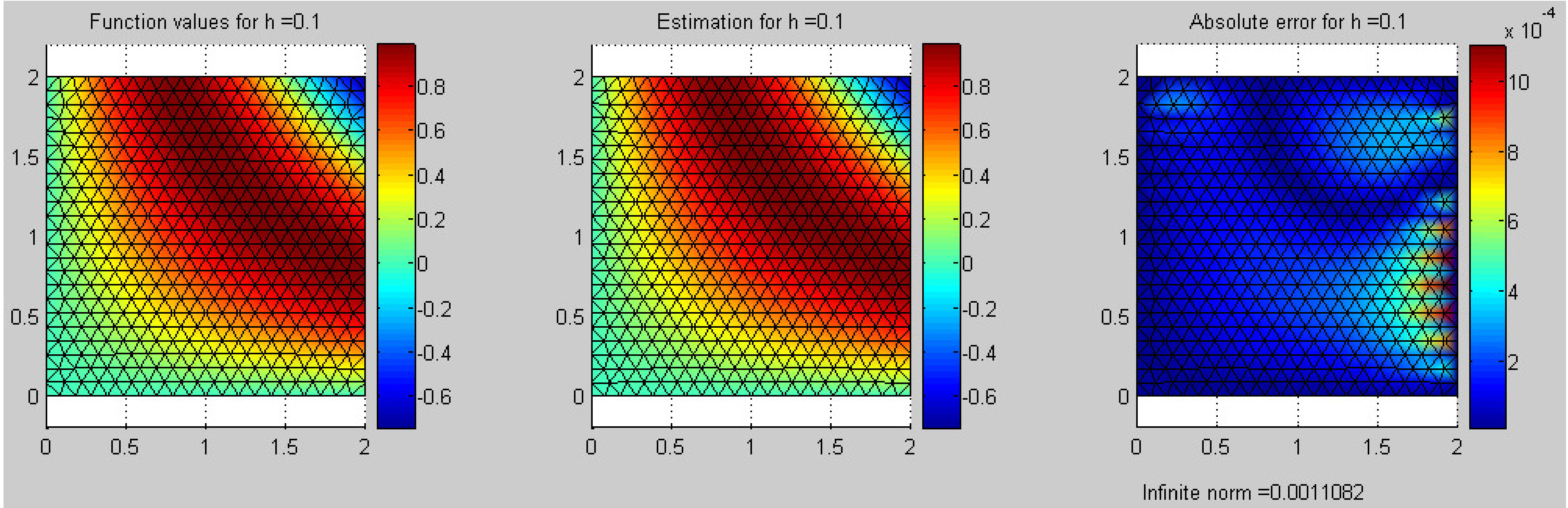
Πειραματικά αποτελέσματα με γραμμικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.4 στο Π2.



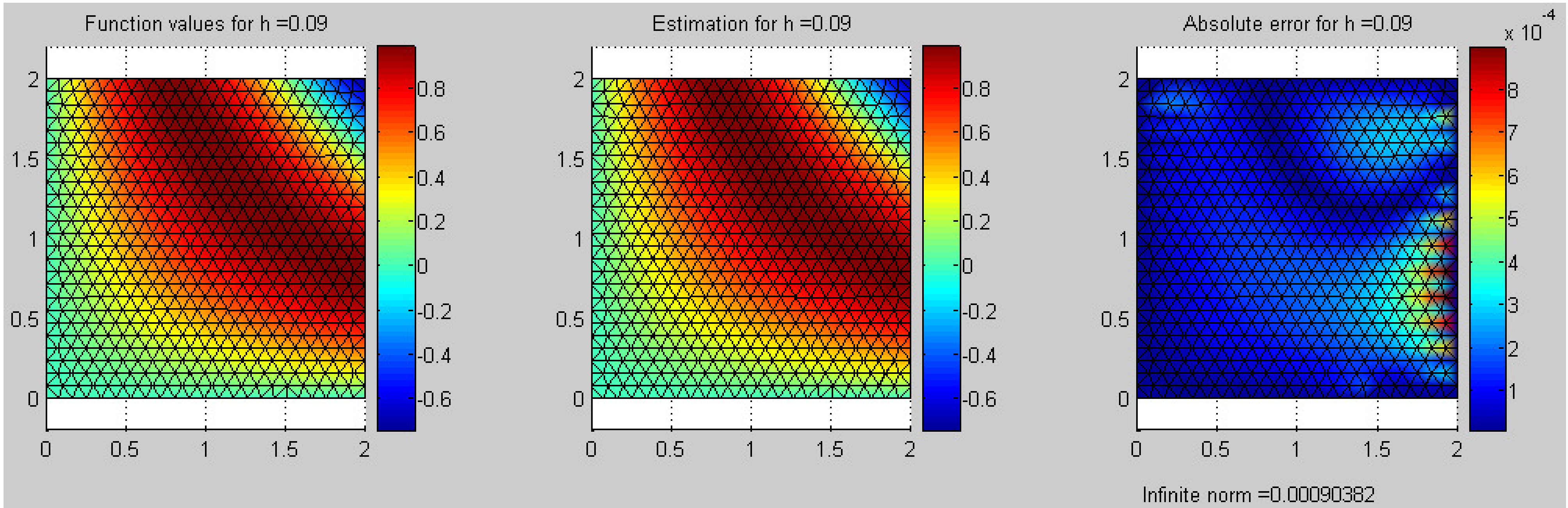
Πειραματικά αποτελέσματα με γραμμικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.3 στο Π2.



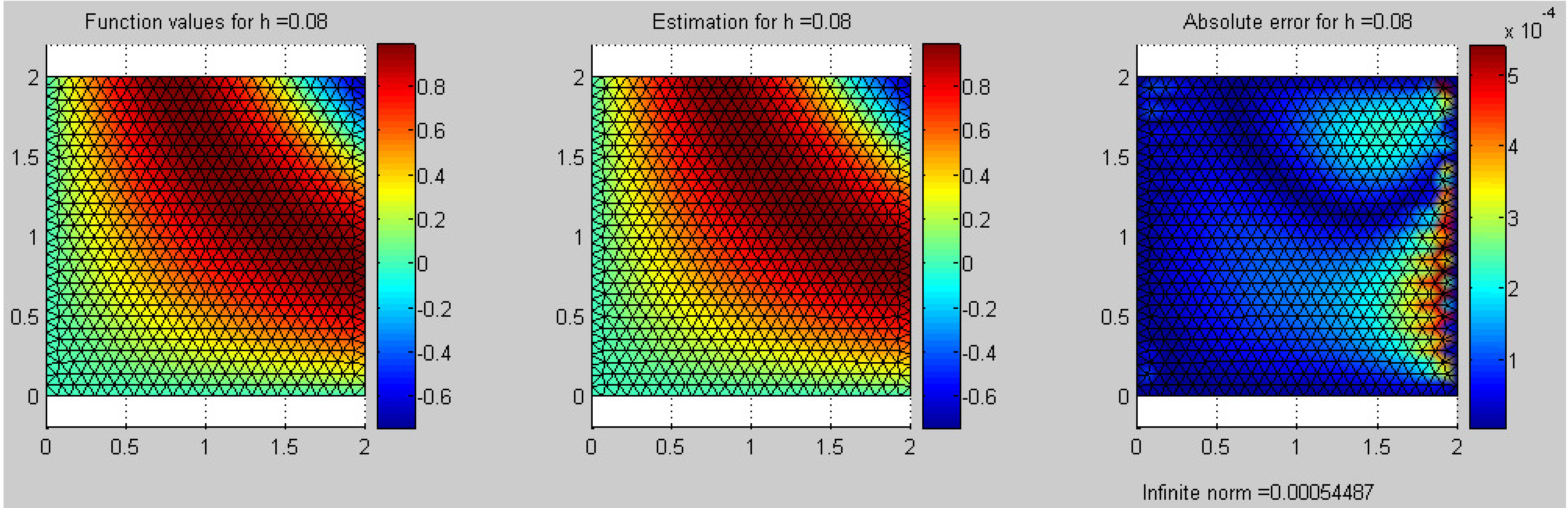
Πειραματικά αποτελέσματα με γραμμικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.2 στο Π2.



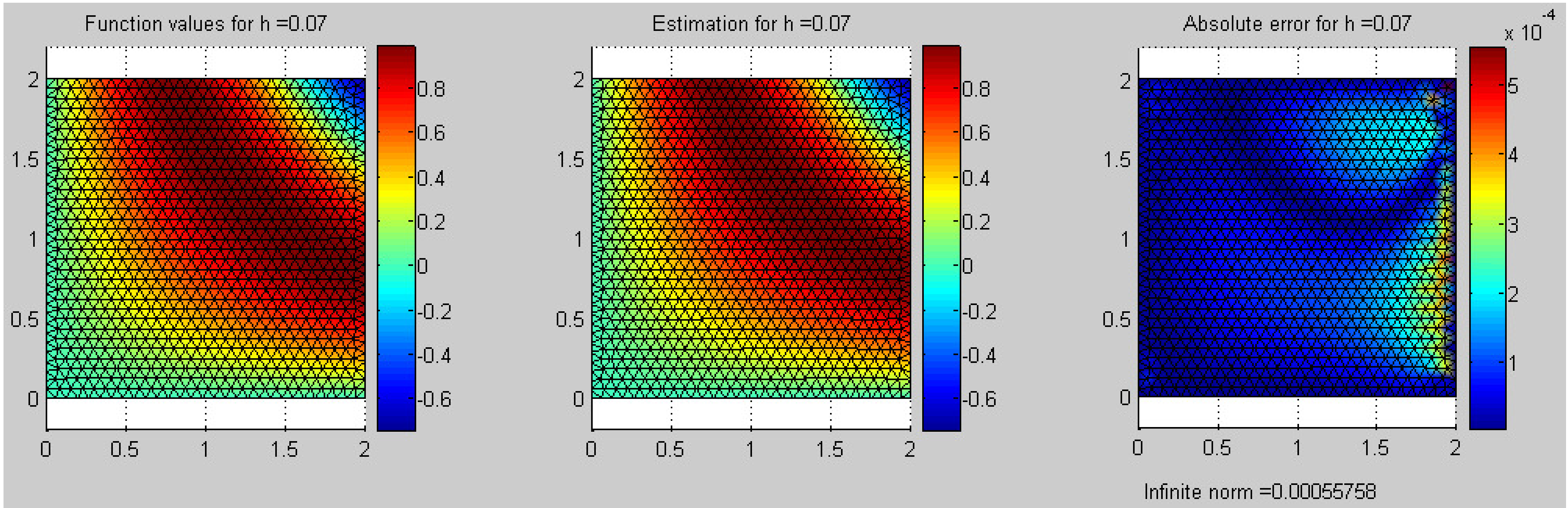
Πειραματικά αποτελέσματα με γραμμικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.1 στο Π2.



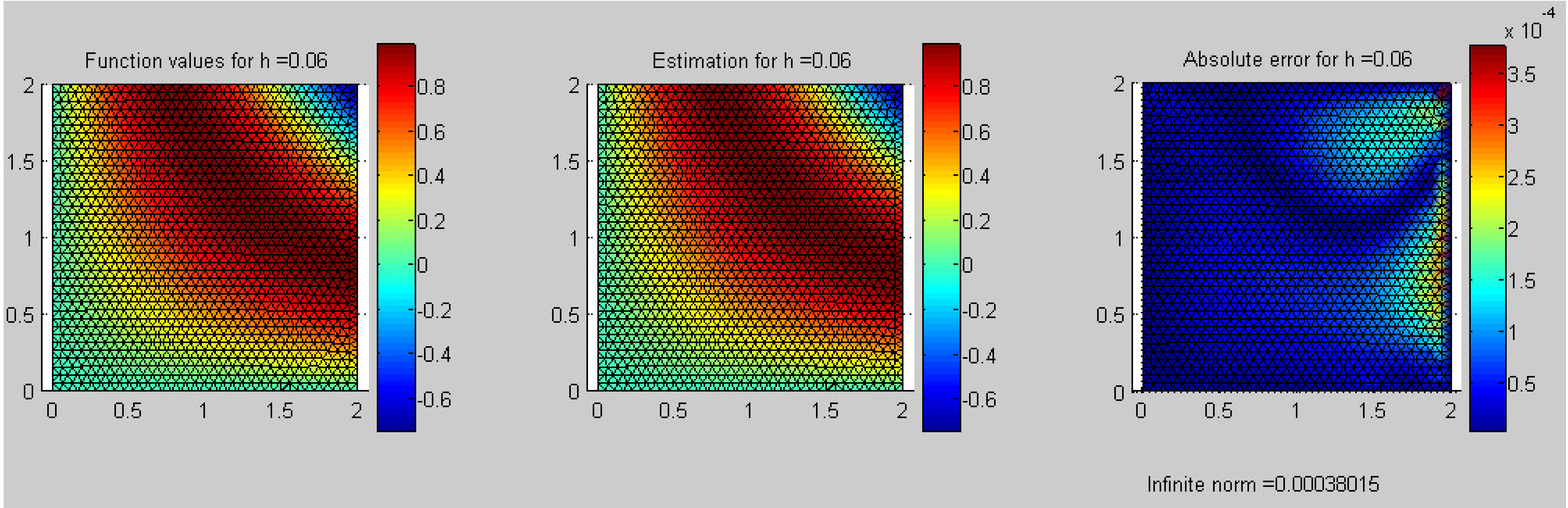
Πειραματικά αποτελέσματα με γραμμικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.09 στο Π2.



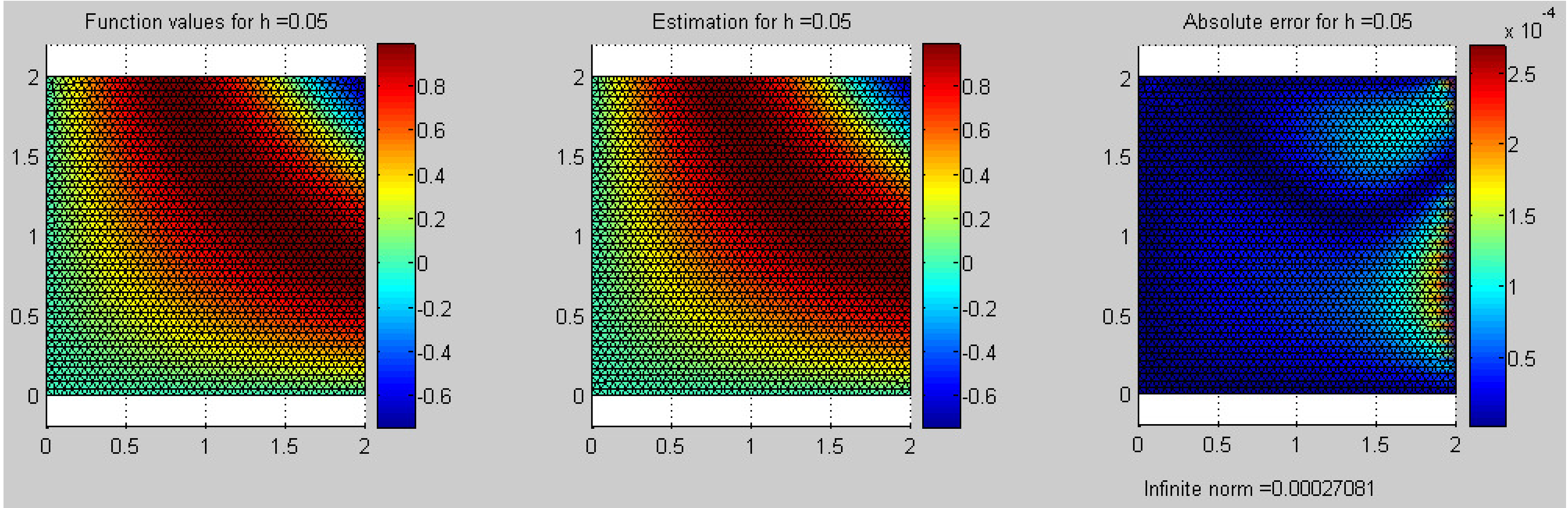
Πειραματικά αποτελέσματα με γραμμικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.08 στο Π2.



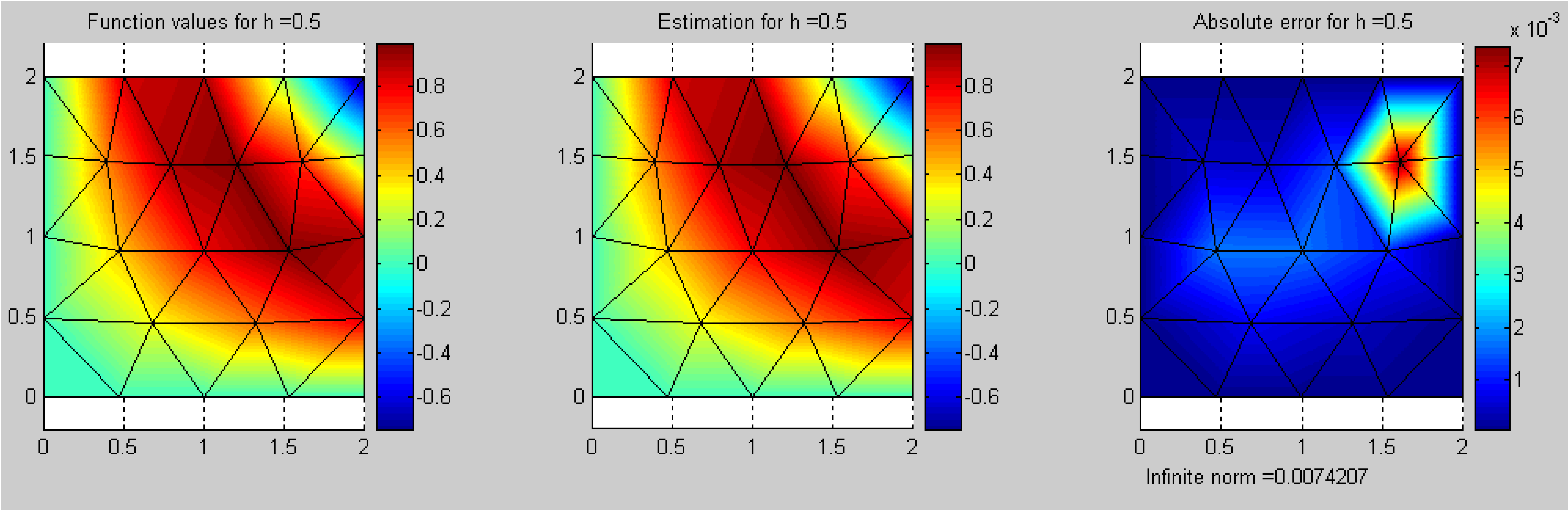
Πειραματικά αποτελέσματα με γραμμικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.07 στο Π2.



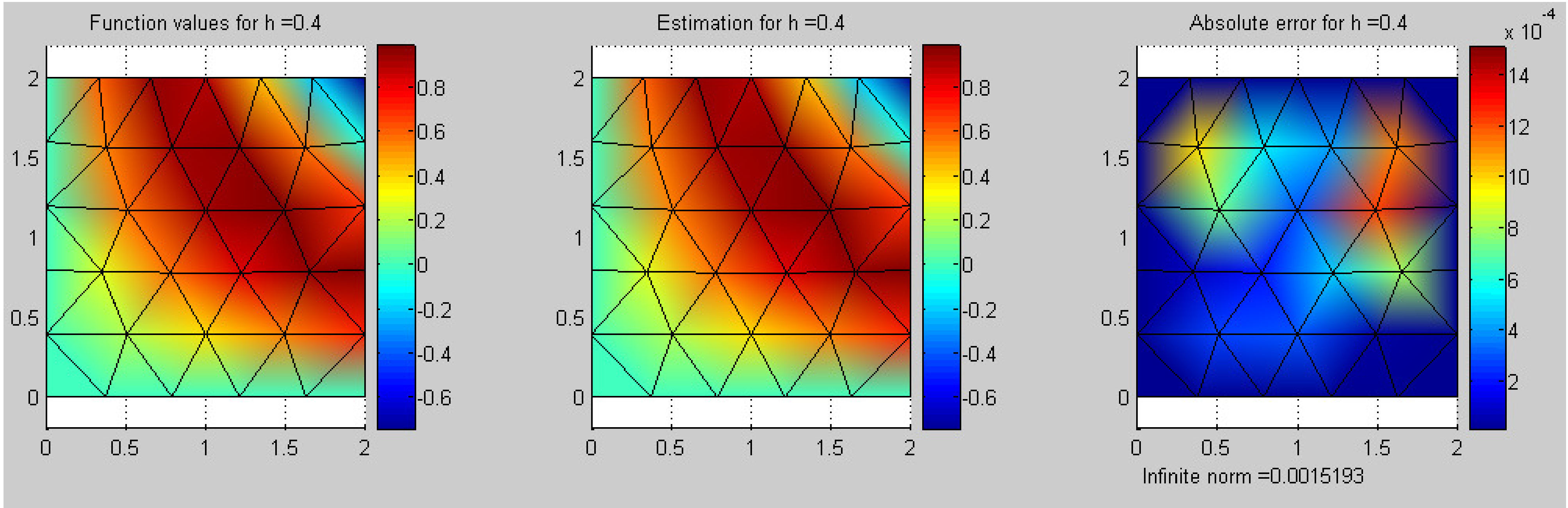
Πειραματικά αποτελέσματα με γραμμικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.06 στο Π2.



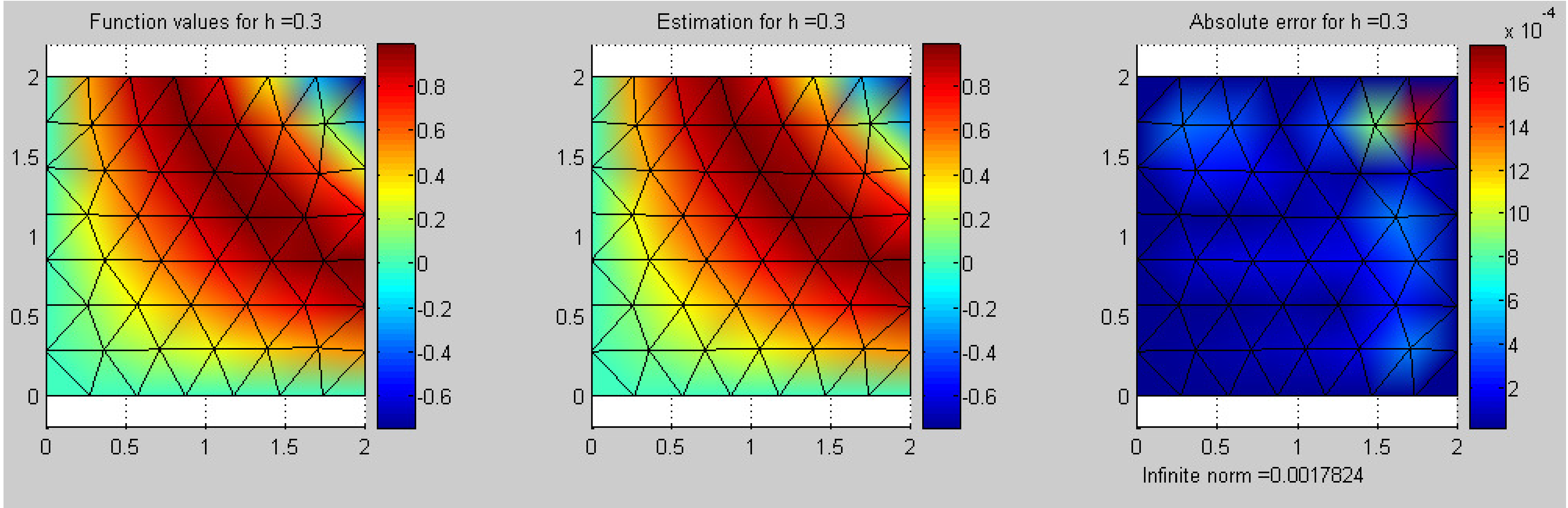
Πειραματικά αποτελέσματα με γραμμικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.05 στο Π2.



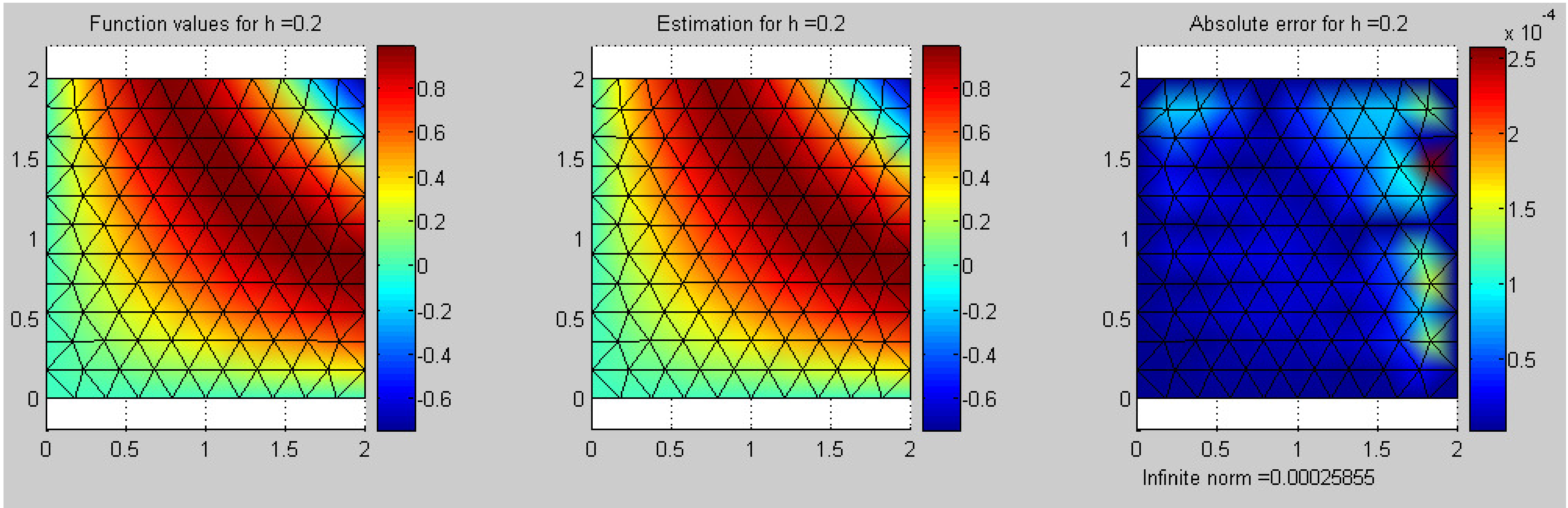
Πειραματικά αποτελέσματα με τετραγωνικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.5 στο Π2.



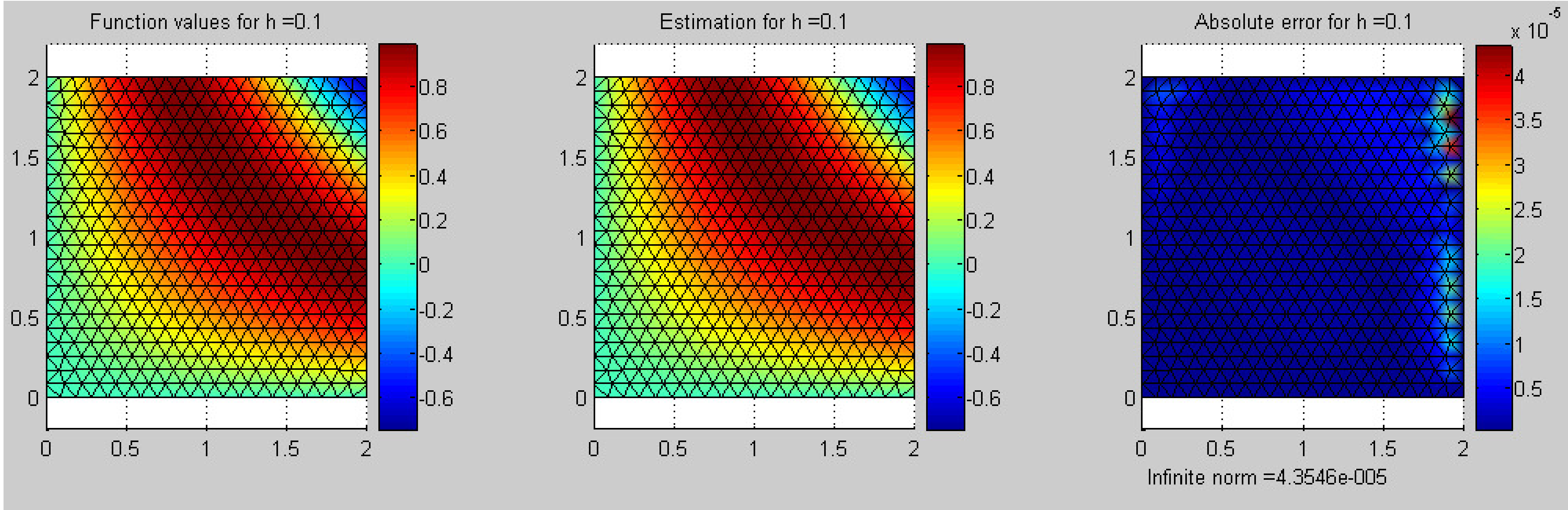
Πειραματικά αποτελέσματα με τετραγωνικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.4 στο Π2.



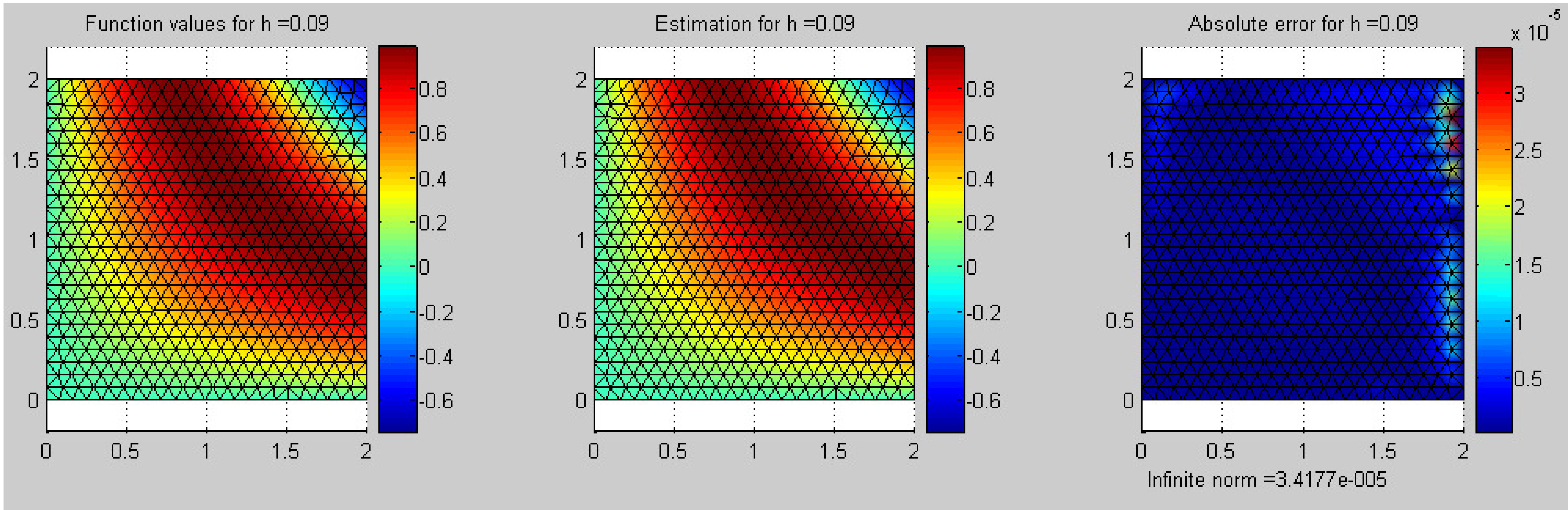
Πειραματικά αποτελέσματα με τετραγωνικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.3 στο Π2.



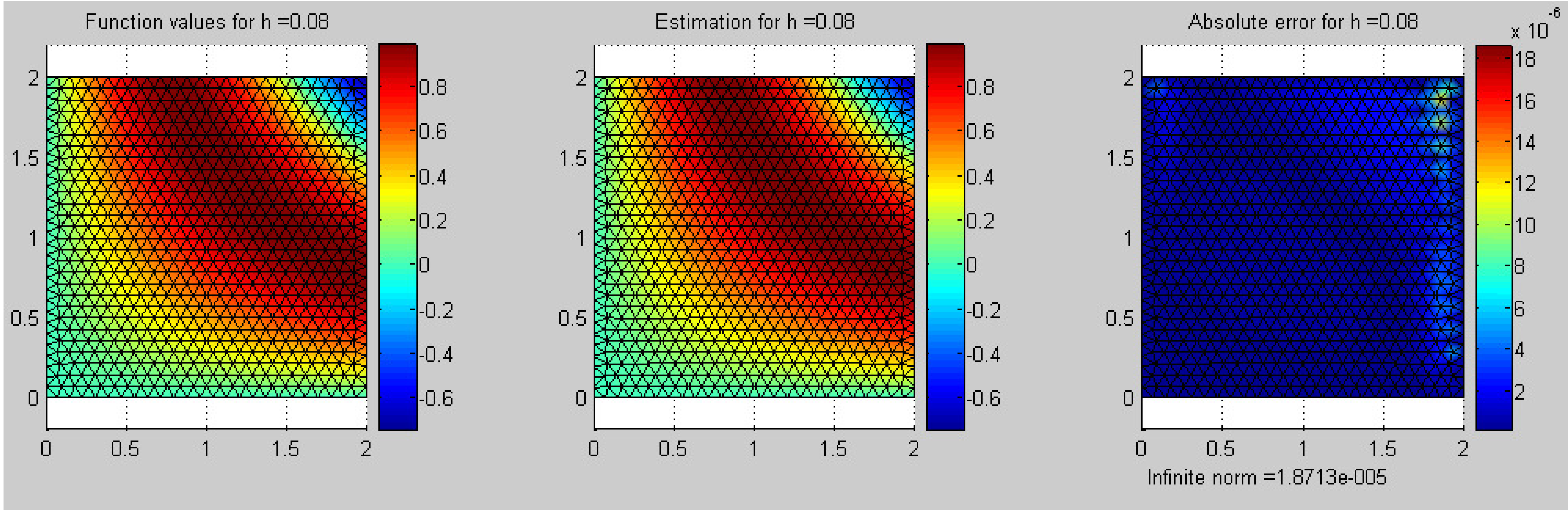
Πειραματικά αποτελέσματα με τετραγωνικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.2 στο Π2.



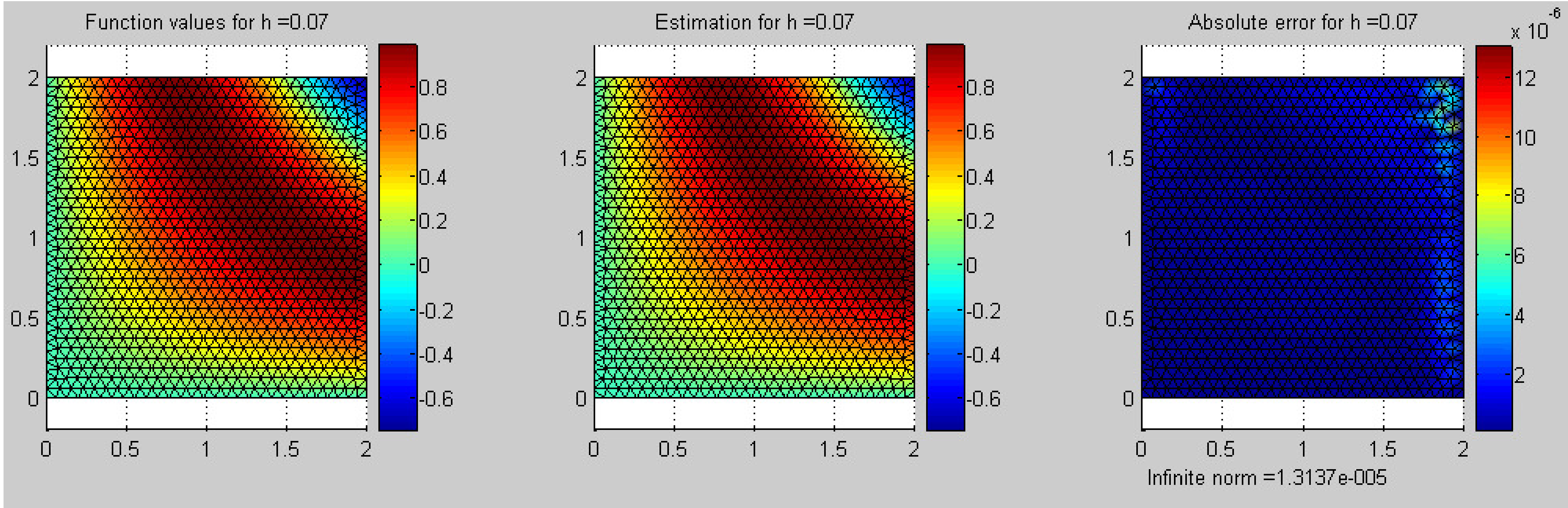
Πειραματικά αποτελέσματα με τετραγωνικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.1 στο Π2.



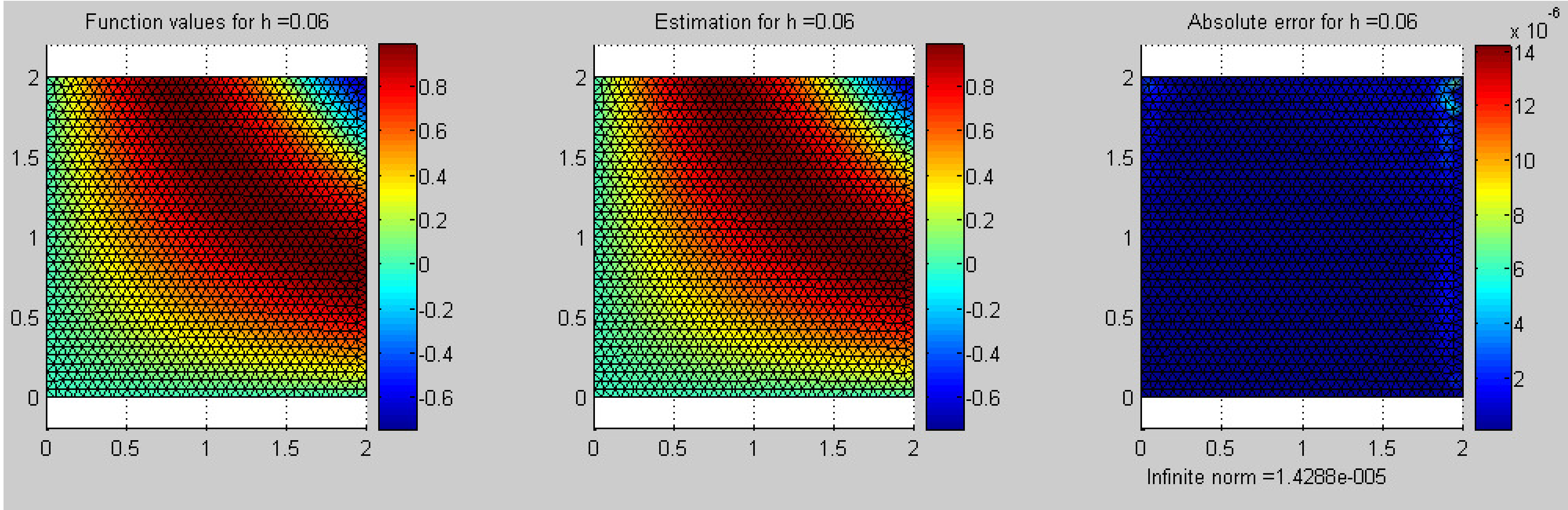
Πειραματικά αποτελέσματα με τετραγωνικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.09 στο Π2.



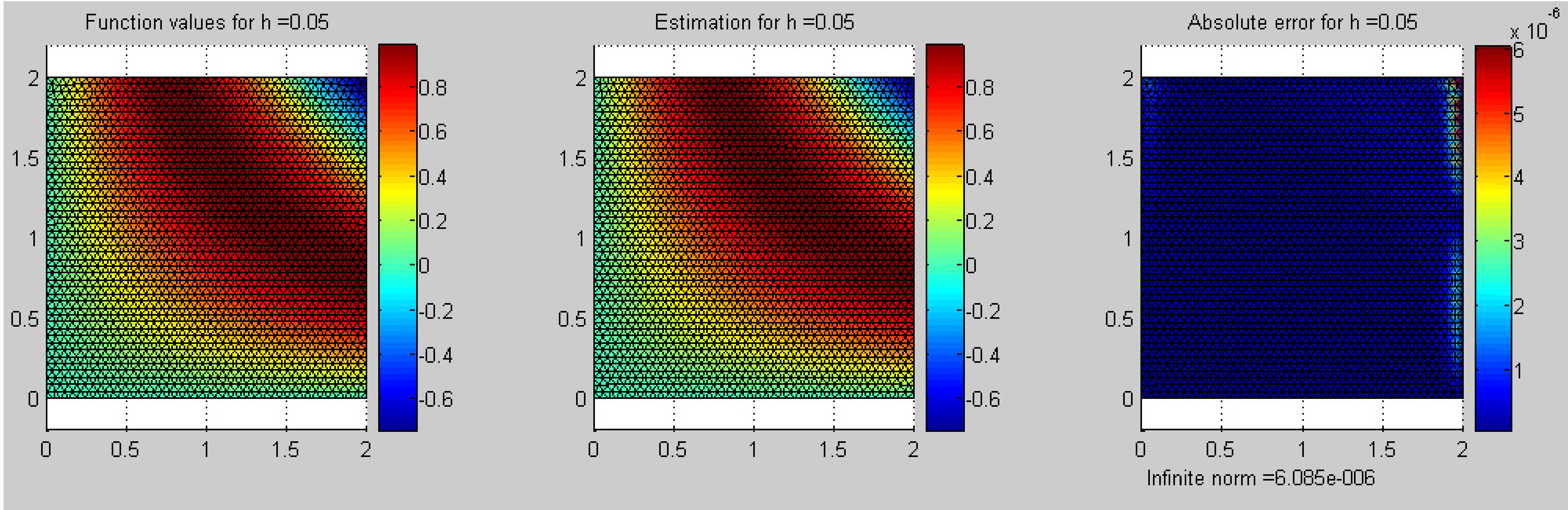
Πειραματικά αποτελέσματα με τετραγωνικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.08 στο Π2.



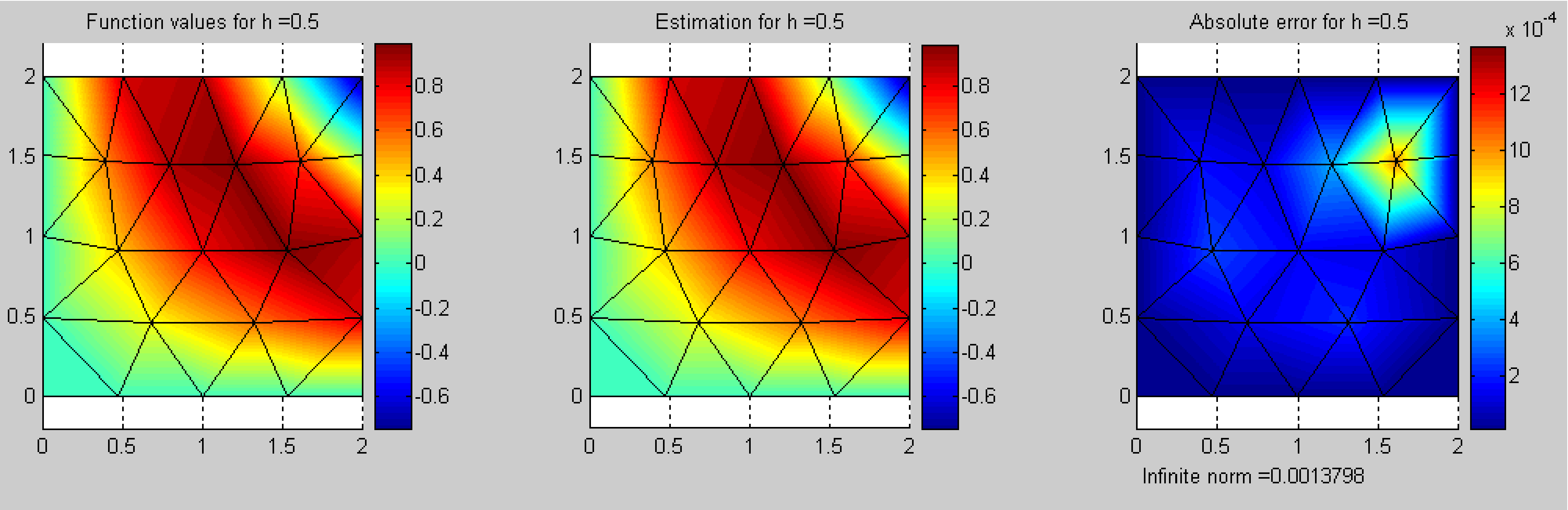
Πειραματικά αποτελέσματα με τετραγωνικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.07 στο Π2.



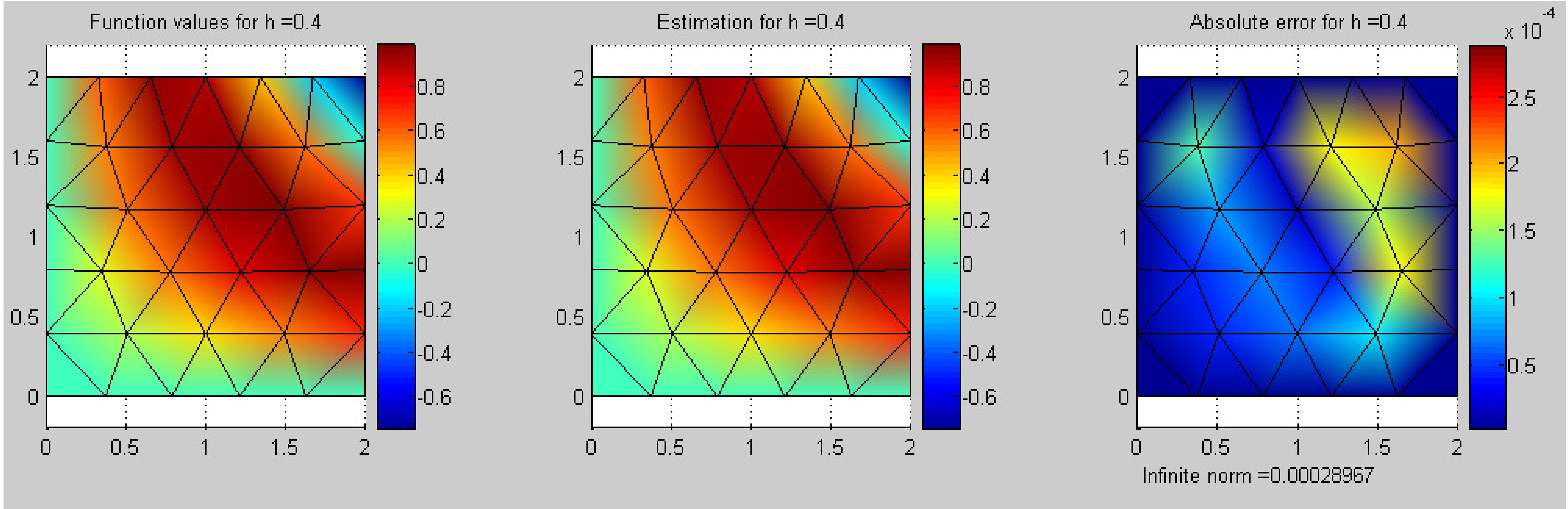
Πειραματικά αποτελέσματα με τετραγωνικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.06 στο Π2.



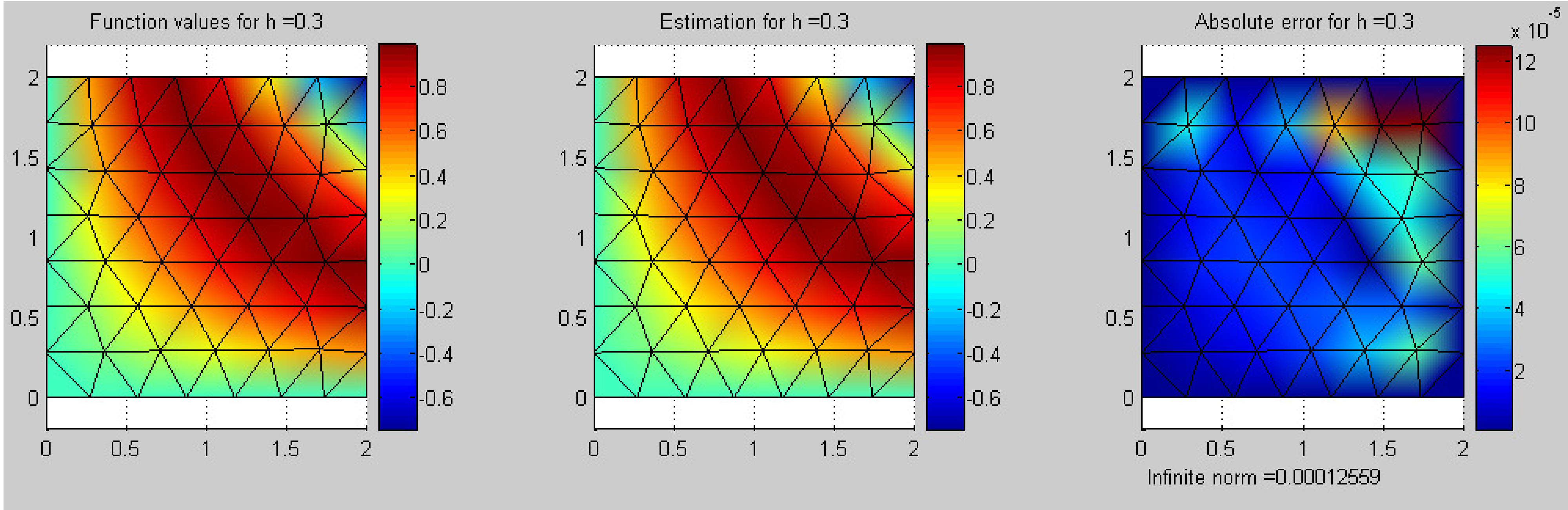
Πειραματικά αποτελέσματα με τετραγωνικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.05 στο Π2.



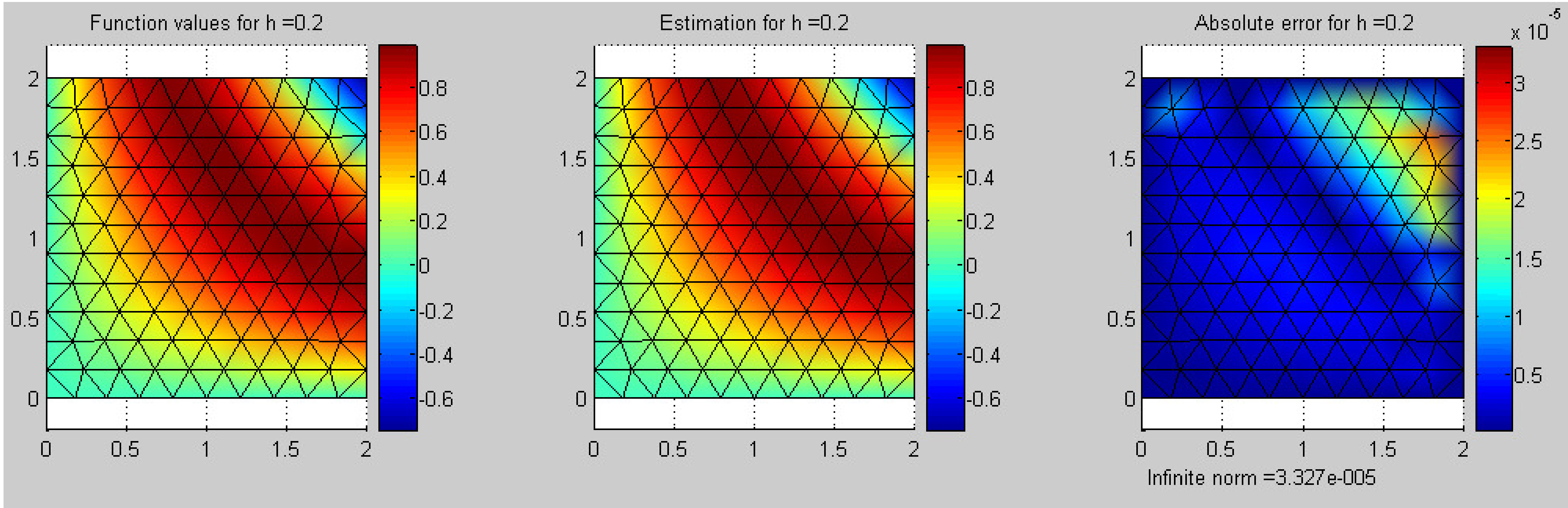
Πειραματικά αποτελέσματα με κυβικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.5 στο Π2.



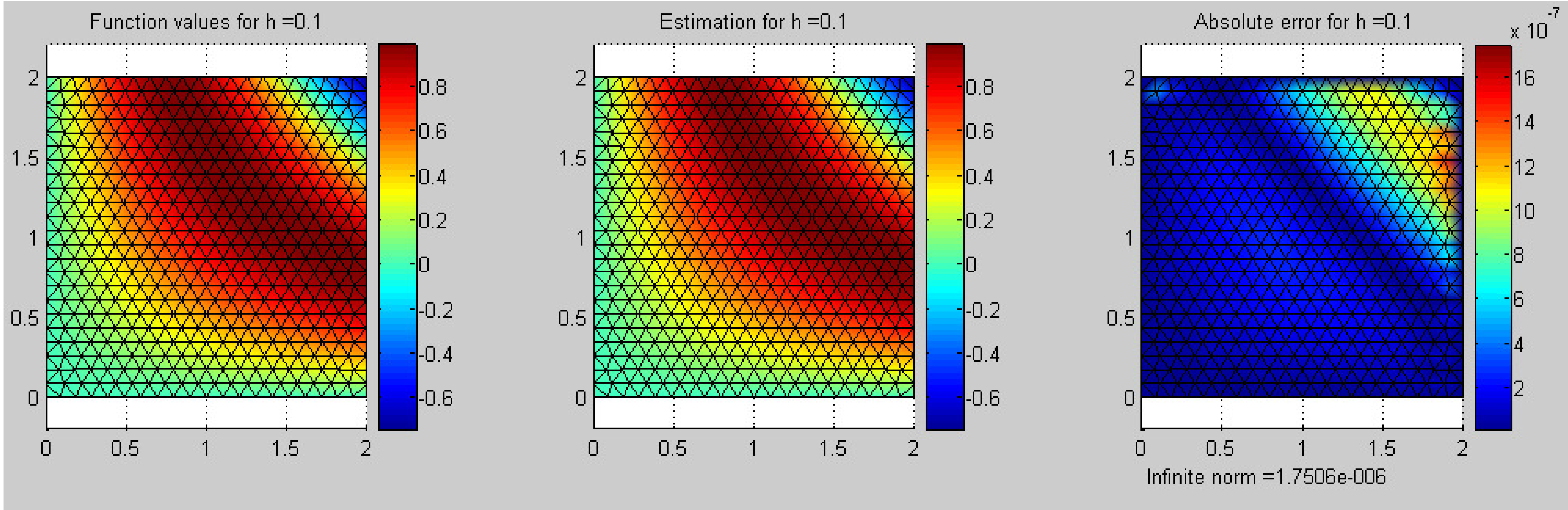
Πειραματικά αποτελέσματα με κυβικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.4 στο Π2.



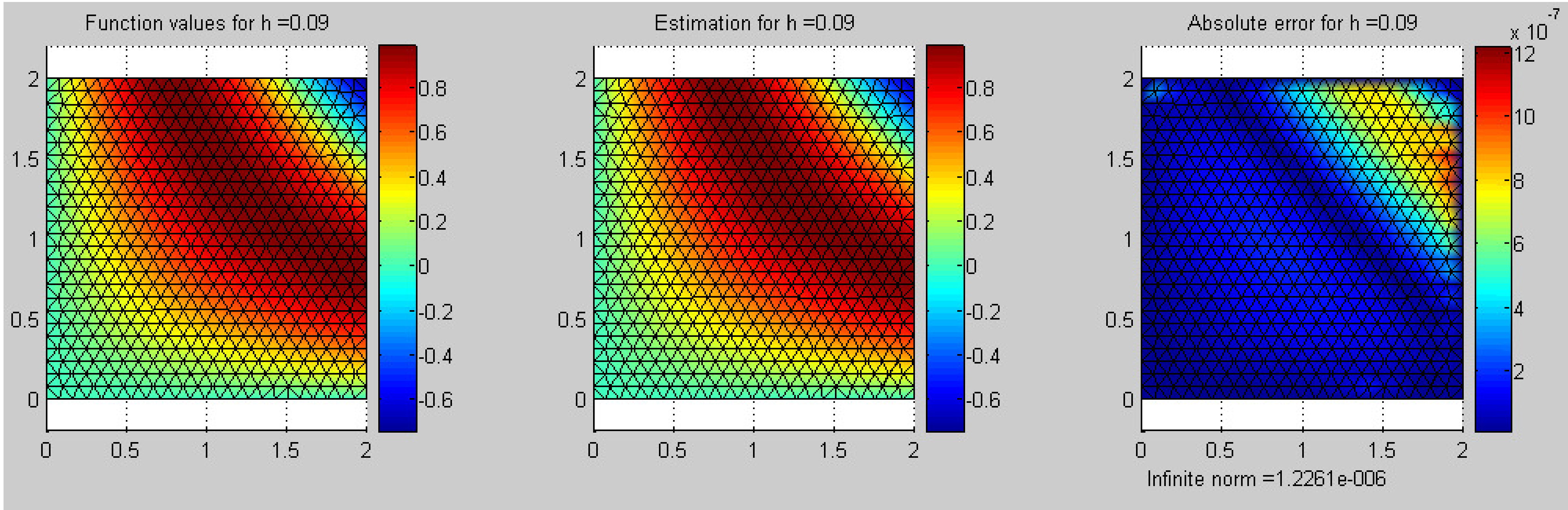
Πειραματικά αποτελέσματα με κυβικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.3 στο Π2.



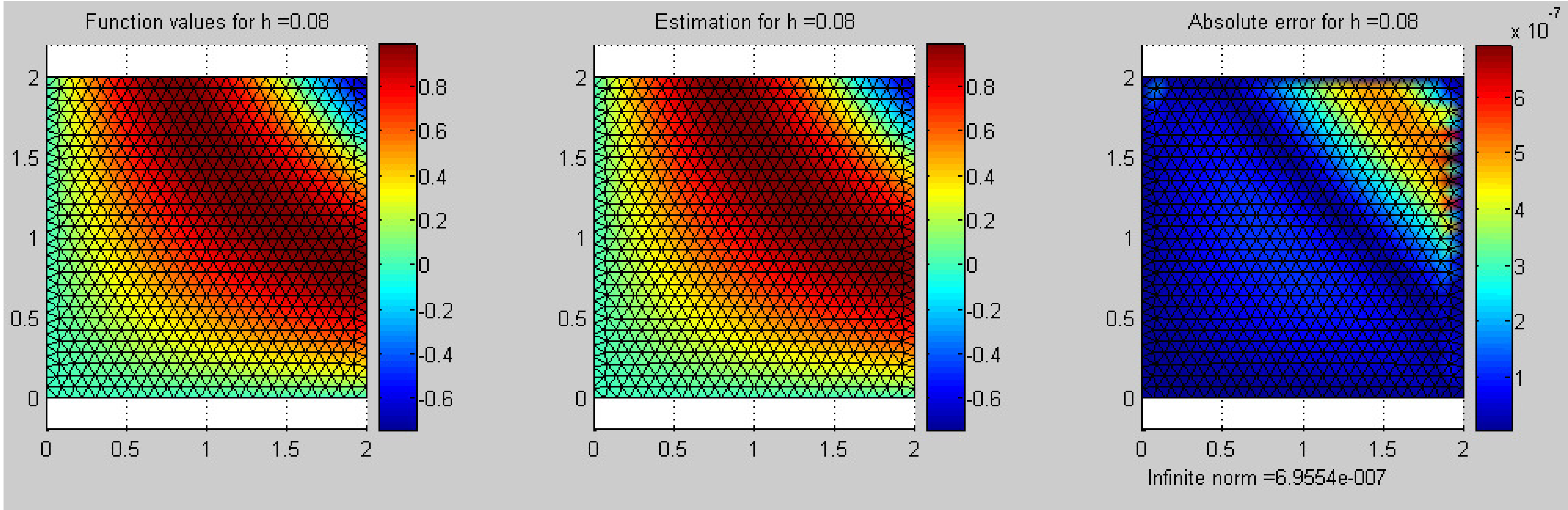
Πειραματικά αποτελέσματα με κυβικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.2 στο Π2.



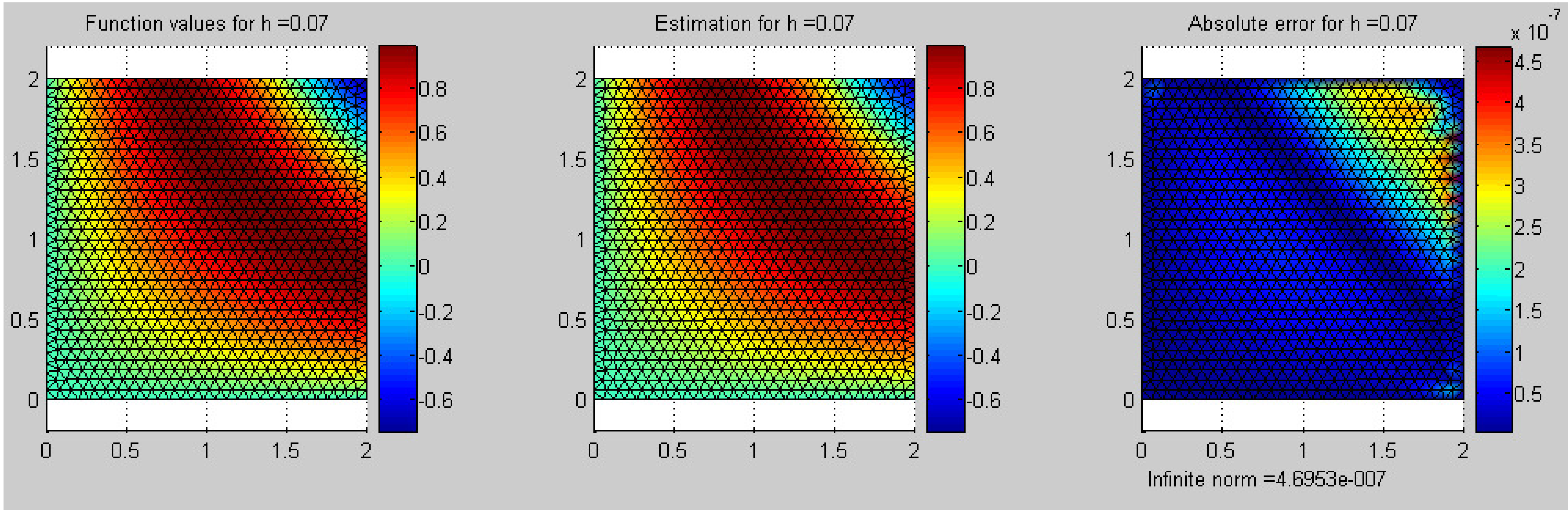
Πειραματικά αποτελέσματα με κυβικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.1 στο Π2.



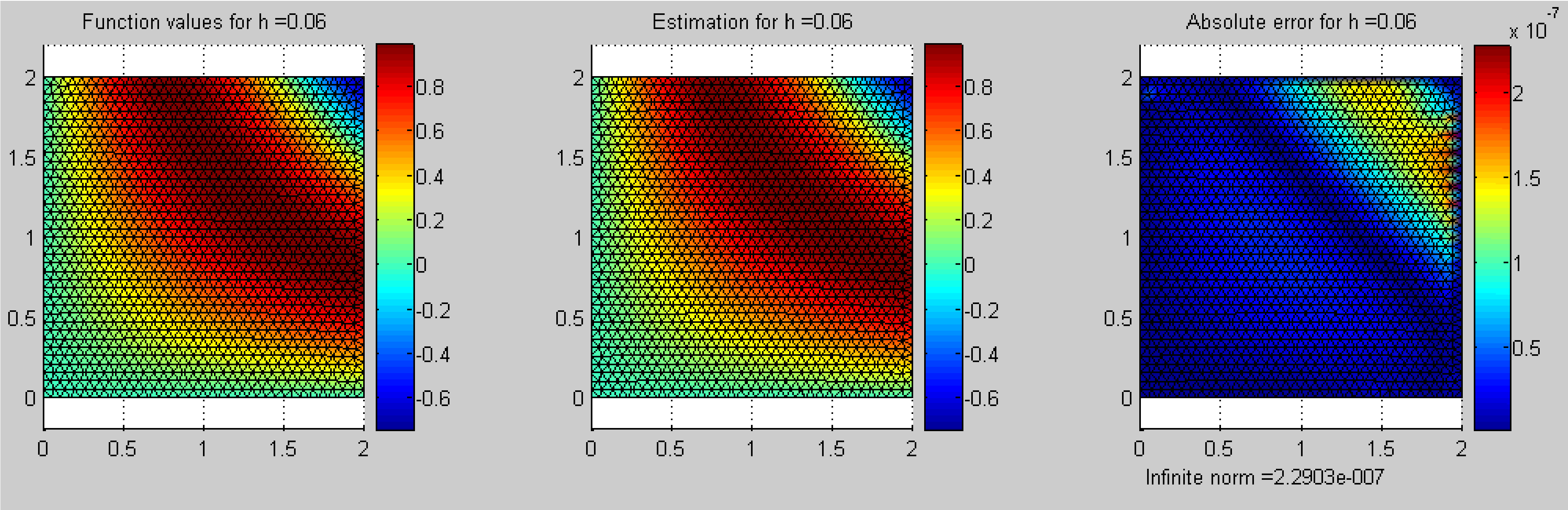
Πειραματικά αποτελέσματα με κυβικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.09 στο Π2.



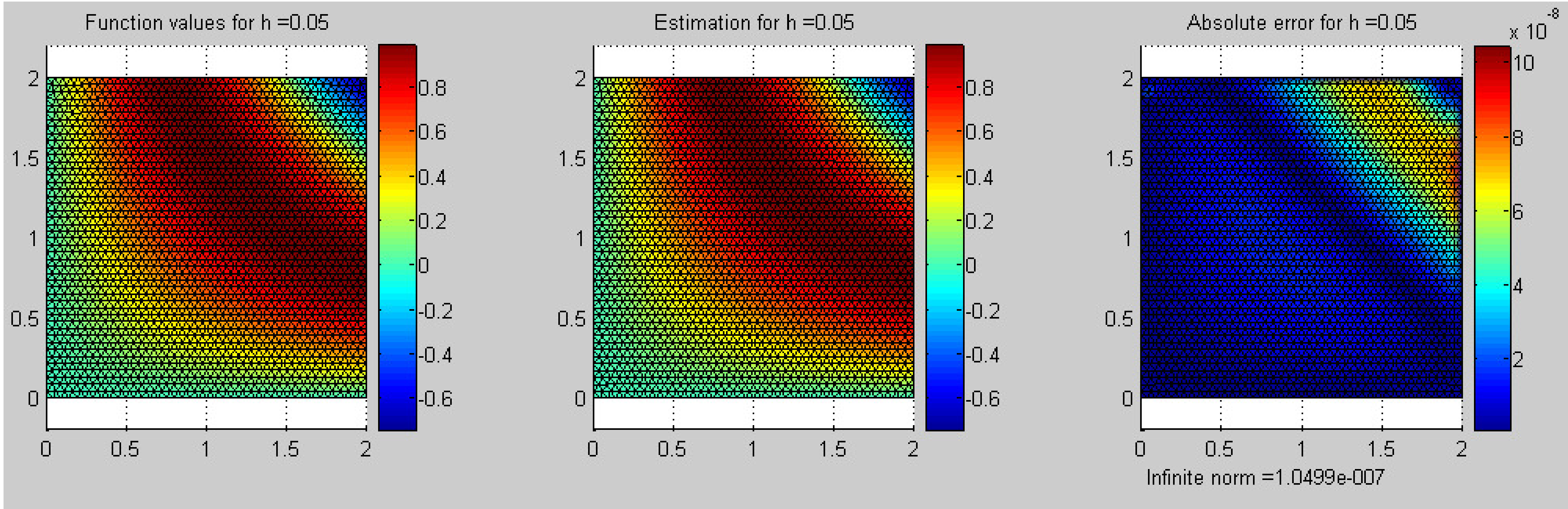
Πειραματικά αποτελέσματα με κυβικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.08 στο Π2.



Πειραματικά αποτελέσματα με κυβικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.07 στο Π2.



Πειραματικά αποτελέσματα με κυβικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.06 στο Π2.



Πειραματικά αποτελέσματα με κυβικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.05 στο Π2.

5. Συμπεράσματα

Μεθοδολογία πειραμάτων

Πρίν παρουσιάσουμε τα συμπεράσματα μας για την αποτελεσματικότητα των μεθόδων, θεωρούμε σκόπιμο να συνοψίσουμε την δομή των πειραμάτων που παραθέσαμε στα αντίστοιχα σημεία των μεθόδων.

Α. Μέθοδος Ritz

Εφαρμογή στις δυο διαστάσεις στην εξίσωση :

$$-\frac{\partial}{\partial x}\left(p(x)\cdot\frac{\partial y}{\partial x}\right)+q(x)\cdot y(x)=f(x,y),\quad p(x)=1,\quad q(x)=\pi^2,\quad f(x,y)=2\pi^2\eta\mu(x)$$

, με αναδρομικό αλγόριθμο έως ότου επιτευχθεί ακρίβεια μικρότερη ή ίση του 10^{-7} με το πλήθος των σημείων να έχει ως αρχική τιμή το δέκα και σε κάθε επανάληψη να αυξάνει κατά δέκα.

Β. Μέθοδος Galerkin

Β.1. Εφαρμογή στη μια διάσταση στην εξίσωση $-u''(x)=f(x),\ 0<x<1,\ u(0)=0,\ u(1)=0$, για τις δώδεκα παρακάτω περιπτώσεις με βήμα 0.01 :

$f_1(x)=1$	$f_2(x)=x$
$f_3(x)=x^2$	$f_4(x)=x^2+x+1$
$f_5(x)=x^3$	$f_6(x)=x^3+x^2+x+1$
$f_7(x)=e^x$	$f_8(x)=e^x\cdot x$
$f_9(x)=e^x\cdot x^2$	$f_{10}(x)=e^x\cdot(x^2+x+1)$
$f_{11}(x)=e^x\cdot(x^3)$	$f_{12}(x)=e^x\cdot(x^3+x^2+x+1)$

Πίνακας 5.1. Συναρτήσεις για την εξίσωση Poisson στη μια διάσταση.

Β.2. Εφαρμογή στις δυο διαστάσεις για τις εξισώσεις :

$$-u_{xx}-u_{yy}=4\quad,\quad -u_{xx}-u_{yy}=y^2\sin(xy)+x^2\sin(xy)$$

, με αναδρομικό αλγόριθμο όπου το βήμα κυμαίνεται από 0.5 έως και 0.05 με ρυθμό μείωσης 0.01 και τα πεδία ορισμού και οι συνοριακές συνθήκες να είναι οι εξής :

	$f(x,y)$	Πεδίο ορισμού Ω	Συνοριακές συνθήκες
Π1	4	$\{(x,y),x^2+y^2=1\}$	$u(x,y)=0,\ (x,y)\in\partial\Omega$
Π2	$y^2\sin(xy)+x^2\sin(xy)$	$[0,2]$ x $[0,2]$	$u(0,y)=u(x,0)=0, u(1,y)=\sin(2y), u(x,1)=\sin(2x)$

Πίνακας 5.2. Σύνολο συνθηκών για την εξίσωση Poisson στις δυο διάστασεις.

Πέραν των πειραμάτων που παρουσιάσαμε παραπάνω εκτελέσαμε περαιτέρω πειράματα με τα κριτήρια τερματισμού να είναι βήμα έως και 0.03 ή ακρίβεια μικρότερη η ίση της τάξης του 10^{-7} χωρίς να παράγουμε τα αντίστοιχα γραφήματα για τα οποία να σημειώσουμε ότι δεν μας προσφέρουν εποπτεία, διότι λόγω της πυκνότητας του πλέγματος δεν είναι διακριτός ο χρωματικός κώδικας. Τα παραπάνω πειράματα υλοποιήθηκαν στην γλώσσα προγραμματισμού Matlab, με τον κώδικα να παρατίθεται με πλήρη σχολιασμό στο παράρτημα και εκτελέστηκαν σε ηλεκτρονικό υπολογιστή με τις εξής προδιαγραφές :

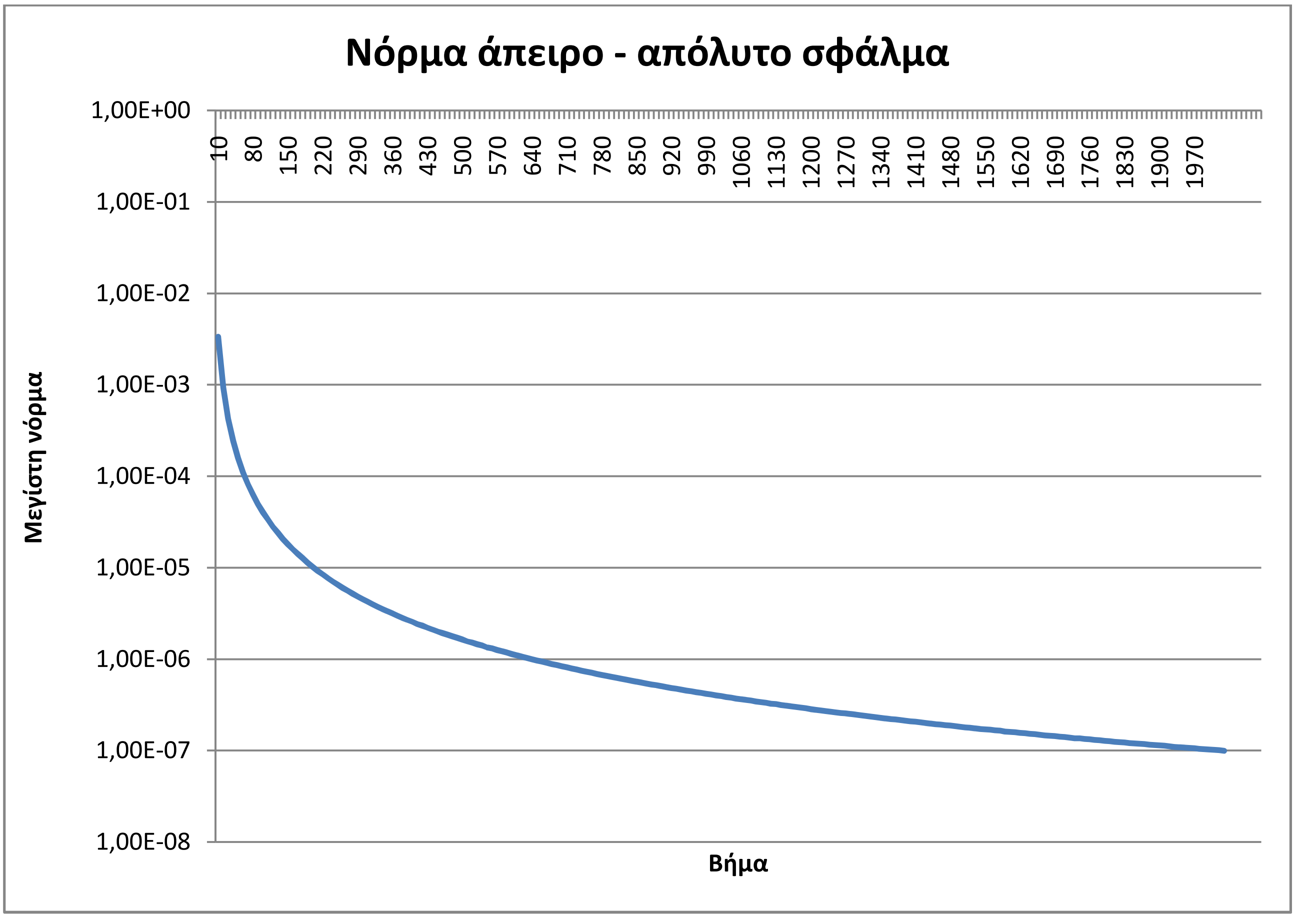
Cpu	**Ram**	**Λειτουργικό σύστημα**	**Έκδοση Matlab**
Pentium 4, 3.2Ghz	512 Mb DDR	Windows Xp, SP3	Version 7.0.1.24704 (R14)

Πίνακας 5.3. Υλικό που χρησιμοποιήθηκε για την διεξαγωγή των πειραμάτων.

Αξιολόγηση αποτελέσματων

Α. Εφαρμογή της μεθόδου Ritz στην εξίσωση $-\frac{\partial}{\partial x}\left(p(x)\cdot\frac{\partial y}{\partial x}\right)+q(x)\cdot y(x)=f(x,y)$

Η μέθοδος Ritz απεδώσε υψηλής ποιότητας αποτελέσματα για $p(x)=1,\;\; q(x)=\pi^2,\;\; f(x,y)=2\pi^2\eta\mu(x)$ όπου διαπιστώθηκε ότι για την νόρμα άπειρο έχουμε ακρίβεια της τάξης του 10^{-8} και στο παρακάτω γράφημα διακρίνουμε την πτωτική πορεία της νόρμας σε λογαριθμική κλίμακα καθώς αυξάνουμε το πλήθος των σημείων :



Γράφημα 5.Α. Τιμές της μεγίστης νόρμας σφάλμα σε λογαριθμική κλίμακα για το απόλυτο σφάλμα.

Από το παραπάνω γράφημα και τα δεδομένα που περιελαμβάνονται στο Ritz.txt προκύπτει ότι :

- Δεν υπάρχουν ταλαντώσεις ή παλινδρομήσεις στην αποτελεσματικότητα του αλγορίθμου, αντιθέτως παρατηρούμε μόνο βελτίωση στην ακρίβεια.
- Επιτυγχάνονται εξαιρετικά επίπεδα για την μεγίστη νόρμα με ακρίβειας της τάξης του 10^{-8} και περαιτέρω προκύπτει ότι :

N = 2030	
Ελάχιστη τιμή	$2{,}00*10^{-10}$
Μέσος όρος σφαλμάτων	$6{,}35*10^{-08}$
Πλήθος σφαλμάτων $< 10^{-9}$	12
Πλήθος σφαλμάτων $< 10^{-8}$	120

Πίνακας 5.4. Στατιστικά για 2030 σημεία.

, συνεπώς υπάρχουν 12 σημεία στα οποία η ακρίβεια ήταν της τάξης 10^{-10}, 108 σημεία όπου η ακρίβεια ήταν της τάξης 10^{-9} και ο μέσος όρος είναι σε εξαιρετικά επίπεδα της τάξης του $6{,}35*10^{-8}$.

- Παρατηρούμε ότι για ακρίβεια μεγίστης νόρμας τάξης 10^{-8} απαιτείται μεγάλο υπολογιστικό κόστος, πιο συγκεκριμένα 2030 σημεία και από την κλίση της καμπύλης ότι έχουμε ραγδαία βελτίωση στην ακρίβεια από 10 έως και 350 περίπου σημεία και μετέπειτα για να υπάρξει αισθητή βελτίωση στην ποιότητα της προσέγγισης χρειαζόμαστε πολύ περισσότερες επαναλήψεις.

Συνοψίζοντας θεωρούμε ότι εάν λάβουμε υπόψη ότι :

- Σφάλματα τάξης 10^{-6} υπολογιστικά είναι ιδιαίτερα ικανοποιητικά [1].
- Τα σφάλματα αποκοπής που επισέρχονται σε υψηλά επίπεδα ακρίβειας σφαλμάτων επηρεάζουν σημαντικά την τελική τιμή της προσέγγισης [2].
- Εκ των πραγμάτων η βελτίωση της προσέγγισης στην κλάση 10^{-6} έως και 10^{-7} είναι δυσκολότερη από ότι στην κλάση 10^{-4} έως και 10^{-5} [3].

, προκύπτει ότι ο αλγόριθμος αποδίδει υψηλής ποιότητας ακρίβεια της τάξης 10^{-6} με μικρό πλήθος απαιτούμενων επαναλήψεων 200 (Πίνακας 3.3.1.Α) και περαιτέρω μας προσφέρει την δυνατότητα για ακόμα καλύτερη προσέγγιση, με εξαιρετικό μέσο όρο της τάξης του 10^{-8} και σημειακά ακρίβεια της τάξης του 10^{-10}, όπου όπως είναι αναμενόμενο μας κοστίζει υπολογιστικά.

Β. Μέθοδος Galerkin

Β.1. Μέθοδος Galerkin στη μια διάσταση

$-u''(x) = f(x),\ 0 < x < 1,\ u(0) = 0,\ u(1) = 0$			
Συνάρτηση	**Απόλυτο σφάλμα**	**Συνάρτηση**	**Απόλυτο σφάλμα**
$f_1(x) = 1$	3.8303e-015	$f_7(x) = e^x$	2.2118e-012
$f_2(x) = x$	1.7070e-015	$f_8(x) = e^x \cdot x$	1.0058e-011
$f_3(x) = x^2$	9.1593e-016	$f_9(x) = e^x \cdot x^2$	3.7124e-011
$f_4(x) = x^2 + x + 1$	6.4115e-015	$f_{10}(x) = e^x \cdot (x^2 + x + 1)$	4.9387e-011
$f_5(x) = x^3$	5.6899e-016	$f_{11}(x) = e^x \cdot (x^3)$	1.0757e-010
$f_6(x) = x^3 + x^2 + x + 1$	6.9111e-015	$f_{12}(x) = e^x \cdot (x^3 + x^2 + x + 1)$	1.5687e-010

Πίνακας 5.5. Πειραματικά αποτελέσματα για 12 περιπτώσεις συναρτήσεων

Στον παραπάνω πίνακα συνοψίσαμε τα αποτελέσματα εφαρμογής της μεθόδου Galerking στην εξίσωση Poisson στη μια διάσταση όπου και λόγω της f_1 επι του πρακτέου αντιμετωπίζουμε και την εξίσωση Laplace στην περίτωση της μια διάστασης, διότι η περίπτωση η συνάρτηση να είναι η μηδενική είναι υπολογιστικά πιο εύκολα αντιμετωπίσιμη. Η μορφή των συναρτήσεων επιλέχθηκε έτσι ώστε να εξετάζουμε τις περιπτώσεις με σταθερή συνάρτηση, με πολυωνυμική συνάρτηση έως και τρίτου βαθμού, εκθετική συνάρτηση αλλά και συνδυασμούς τους. Από τα παραπάνω αποτελέσματα προκύπτει ότι η ακρίβεια της προσέγγισης είναι εξαιρετική με την χειρότερη περίπτωση να είναι αυτή όπου το μέγιστο σφάλμα είναι της τάξης του 10^{-10} για την f_{12} , όπου με περαιτέρω ανάλυση προκύπτει ότι :

N = 101	
Μέγιστη τιμή	1,56E-10
Ελάχιστη τιμή	4,58E-12
Μέσος όρος σφαλμάτων	1,02E-10
Πλήθος σφαλμάτων $< 10^{-11}$	5
Πλήθος σφαλμάτων $< 10^{-10}$	37
Πλήθος σφαλμάτων $< 10^{-9}$	59

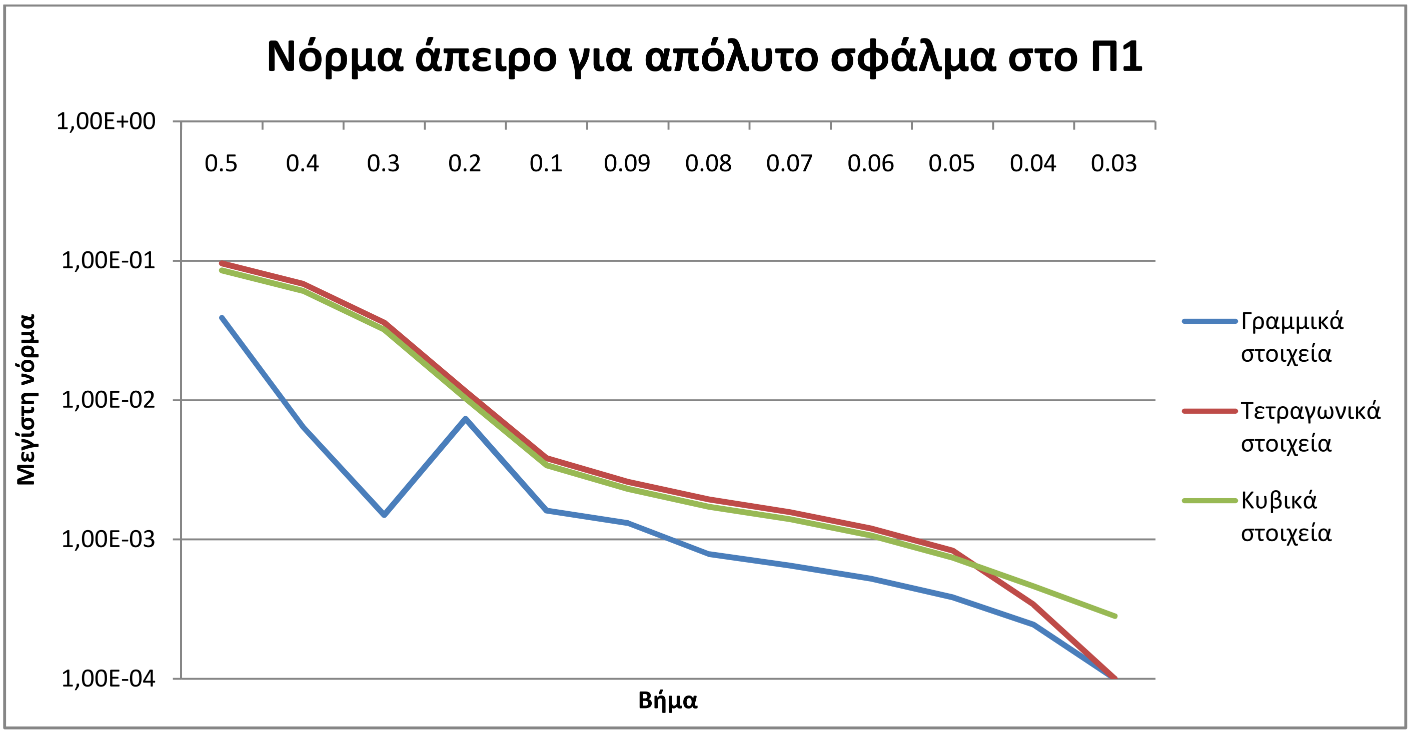
Πίνακας 5.6. Στατιστικά για 101 σημεία.

, συνεπώς υπάρχουν 5 σημεία στα οποία η ακρίβεια ήταν της τάξης 10^{-12}, 42 σημεία όπου η ακρίβεια ήταν της τάξης 10^{-11} ή ισοδύναμα ακόμα και στην χειρότερη περίπτωση απόδοσης όπου έχουμε μέγιστη τιμή σφάλματος 10^{-10} το 42/29=41,58% των τιμών των σφαλμάτων είναι μικρότερες της τάξης του 10^{-10} και ο μέσος όρος είναι σε εξαιρετικά επίπεδα της τάξης του $4{,}58 \cdot 10^{-10}$.

B.2. Μέθοδος Galerkin στις δυο διαστάσεις

Π1	$-u_{xx}-u_{yy}=4$		
	Νόρμα άπειρο		
h	**Γραμμικά στοιχεία**	**Τετραγωνικά στοιχεία**	**Κυβικά στοιχεία**
0.5	3,90E-02	9,59E-02	8,52E-02
0.4	6,39E-03	6,83E-02	6,07E-02
0.3	1,50E-03	3,60E-02	3,20E-02
0.2	7,34E-03	1,16E-02	1,03E-02
0.1	1,61E-03	3,84E-03	3,41E-03
0.09	1,31E-03	2,60E-03	2,31E-03
0.08	7,86E-04	1,94E-03	1,72E-03
0.07	6,48E-04	1,57E-03	1,40E-03
0.06	5,23E-04	1,20E-03	1,07E-03
0.05	3,85E-04	8,32E-04	7,40E-04
0.04	2,44E-004	5,18 E-04	4,60E-04
0.03	1.26E-004	3,41E-04	2,82E-04

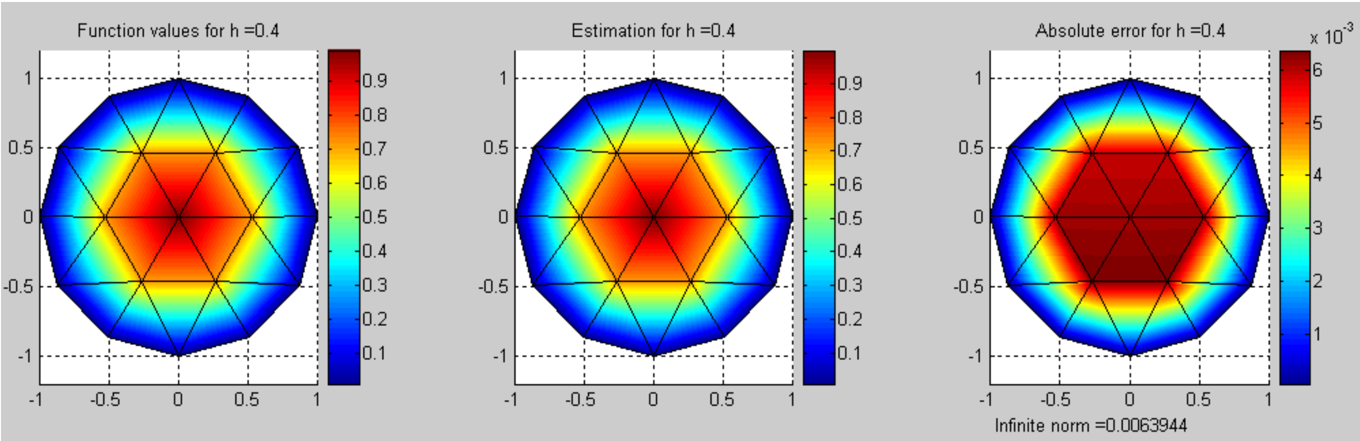
Πίνακας 5.7. Τιμές της νόρμα άπειρο στο Π1, για διαφορετικούς τύπους στοιχείων με βήμα έως και 0.03 .



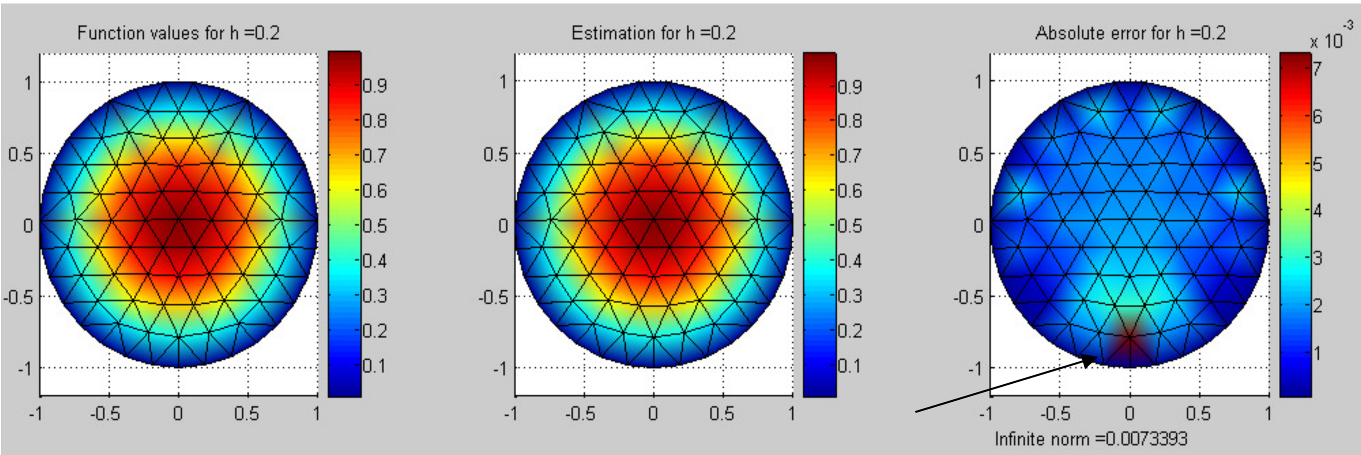
Γράφημα 5.Β. Μεγίστη νόρμα σε λογαριθμική κλίμακα, με τρείς τύπους και διάφορες τιμές εκλέπτυνσης στο Π1.

Το πρόβλημα Π1 εμφανίζει ιδιαίτερες δυσκολίες διότι προσπαθούμε με πλέγμα τριγωνικής μορφής να προσεγγίσουμε συνοριακή μορφή η οποία είναι κυκλική, συνεπώς αναφορικά με την μεγίστη νόρμα είναι αναμενόμενο ότι θα υπάρχουν τιμές για τις οποίες θα υπάρχει απόκλιση και εάν συνυπολογίσουμε ότι για την μεγίστη νόρμα αρκεί μια τέτοια περίπτωση έτσι ώστε να αυτή να εκτιναχθεί αριθμητικά αναμένουμε ταλαντώσεις.

Στο σημείο αυτό για την μεγίστη νόρμα θέλουμε να σημειώσουμε ότι μας προσφέρει ιδιαίτερη ευκολία στον υπολογισμό ειδικά εάν συνυπολογίσουμε το υπολογιστικό κόστος άλλων μετρικών για μεγάλο πλήθος σημείων και ότι εννοιολογικά μας προσφέρει πολύ ισχυρά αποτελέσματα διότι εάν το μέγιστο σφάλμα είναι υψηλής ακρίβειας έχουμε εξαιρετική απόδοση. Παρόλα ταύτα έχει την αδυναμία ότι αρκεί να παρουσιαστεί σημειακά μεγάλη απόκλιση για να λάβουμε παραπλανητικά αποτελέσματα, συνεπώς για την αξιολόγηση της αποδοτικότητας του αλγορίθμου χρειάζεται είτε να ανατρέξουμε στα αποτελέσματα για όλα τα σημεία είτε να παράξουμε στατιστικά απ’ αυτά είτε όπως θεωρούμε ως καλύτερη λύση να χρησιμοποιήσουμε χρωματικό κώδικα, π.χ. για την περίπτωση των γραμμικών στοιχείων με h=0.3 και h=0.2 έχουμε ότι η μεγίστη νόρμα είναι $6{,}39*10^{-03}$ και $7{,}34*10^{-03}$ αντίστοιχα, συνεπώς μπορεί εκ πρώτης να καταλήξουμε στο ότι υπάρχει ταλάντωση στην συμπεριφορά του αλγορίθμου.



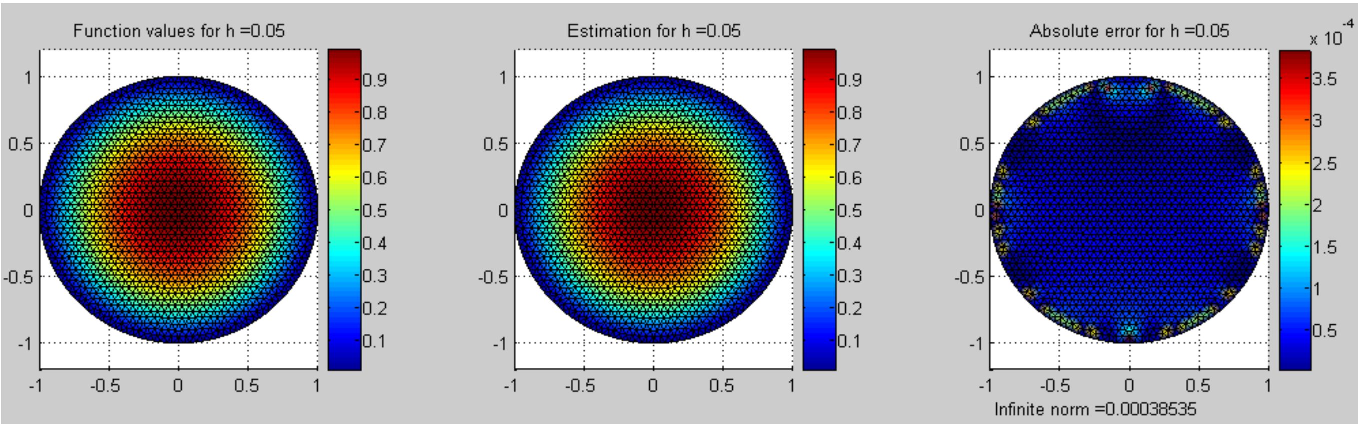
Γραμμικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.3 στο Π1.



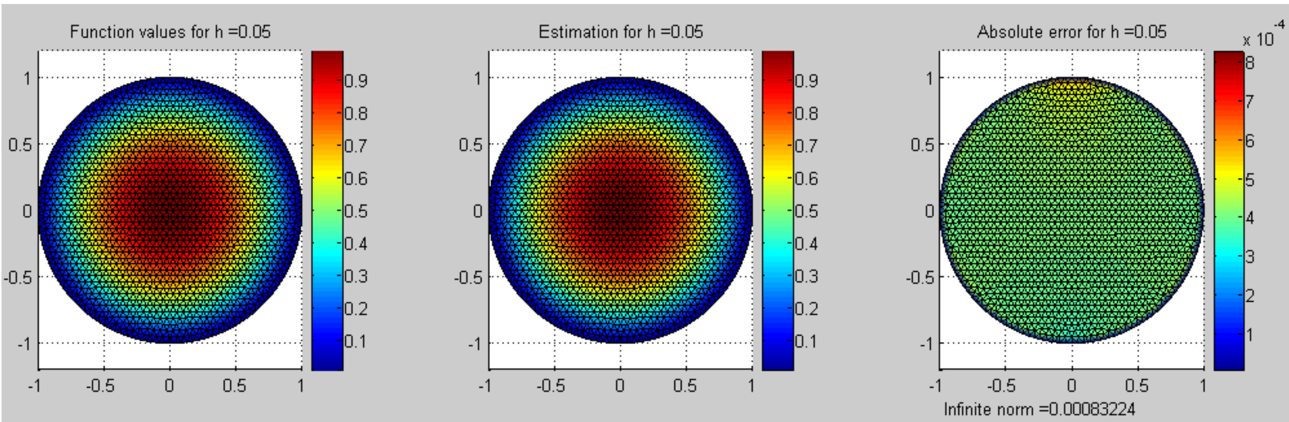
Γραμμικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.2 στο Π1.

Από τον χρωματικό κώδικα όμως γίνεται ξεκάθαρο ότι η μέση αποτελεσματικότητα του αλγορίθμου για την περίπτωση όπου h=0.2 είναι σαφώς καλύτερη, διότι το κυανό ή ισοδύναμα χαμηλές τιμές της τάξης 10^{-3} καλύπτουν σχεδόν ολόκληρο τον δίσκο σε αντίθεση με h=0.4 όπου κυριαρχούν τιμές προς το ερυθρό, συνεπώς η αριθμητική μας προσέγγιση είναι πολύ καλύτερη κατά μέσο όρο αλλά λόγω της υφής του συνόρου υπήρξε ένα σημείο όπου η προσέγγιση ήταν χειρότερη (Σημειώνεται με βέλος στο αντίστοιχο γράφημα), εξ' ου και τα αποτελέσματα για την μεγίστη νόρμα.

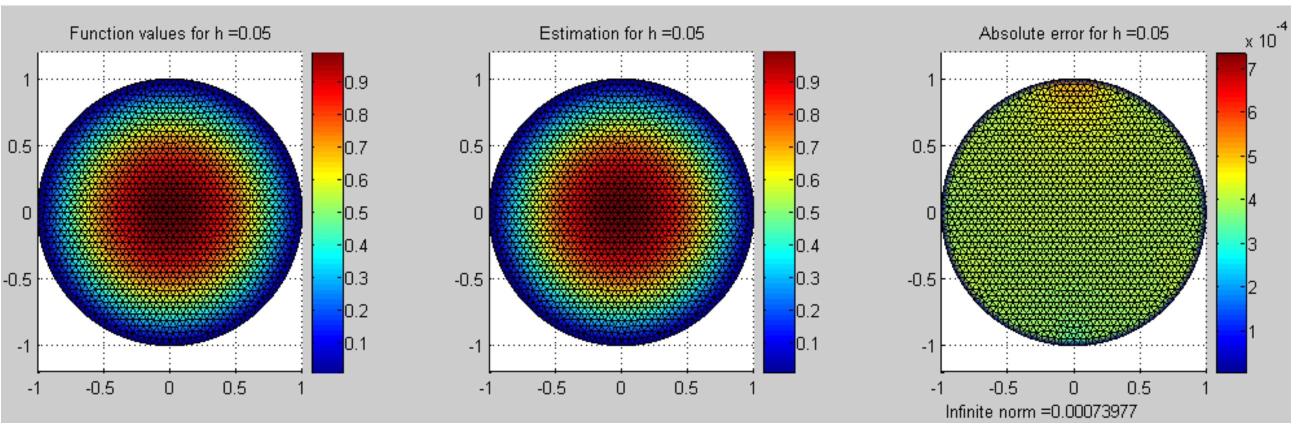
Η παραπάνω παρατήρηση εκτείνεται και στο να μας παράσχει εξήγηση στην σύγκριση των μη αναμενόμενων αποτελεσμάτων ανάμεσα στους τύπους των στοιχείων, όπου η λογική σειρά θα ήταν τα βέλτιστα αποτελέσματα να μας παρουσιαζόταν από τα κυβικά στοιχεία, ακολουθούμενα από τα τετραγωνικά και τέλος από τα γραμμικά. Θεωρούμε ότι η σωστή αξιολόγηση της αποτελεσμάτικοτητας του αλγορίθμου ειδικά όταν οι διαφορές στην μεγίστη νόρμα είναι είναι ιδιαίτερα μικρές βασίζεται στην απόδοση τους σε όλο το πεδίο ορισμού, π.χ. για h=0.05 οι τιμές της μεγίστης νόρμας είναι $3{,}85 \cdot 10^{-04}$, $8{,}32 \cdot 10^{-04}$ και $7{,}40 \cdot 10^{-04}$ για τα γραμμικά, τα τετραγωνικά και τα κυβικά στοιχεία αντίστοιχα.



Γραμμικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.05 στο Π1.



Τετραγωνικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.05 στο Π1.



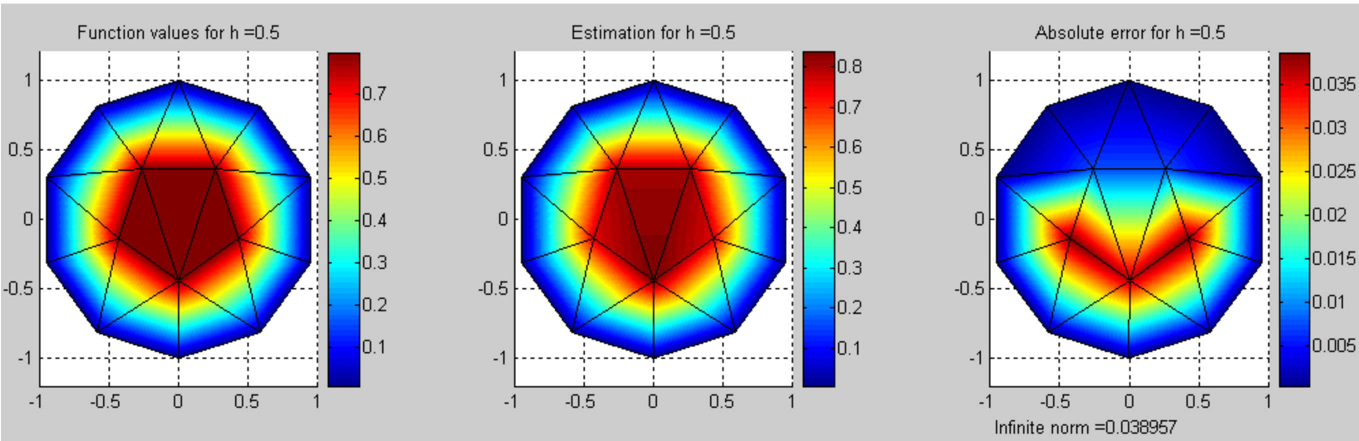
Κυβικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.05 στο Π1.

Από τον χρωματικό κώδικα παρατηρούμε ότι για τα γραμμικά στοιχεία έχουμε κυανό χρώμα εντός του δίσκου αλλά πολλαπλά σημεία προς το ερυθρό καθώς εξετάζουμε το σύνορο σε αντίθεση με τα τετραγωνικά και τα κυβικά στοιχεία όπου εντός του δίσκου εμφανίζονται ελαφρώς χειρότερα αποτελέσματα αλλά κυανό χρώμα στην συνοριακή περιοχή. Έως εκ τούτου προκύπτουν τα εξής:

- Επί του συνόρου τα κυβικά και τα τετραγωνικά στοιχεία επιτυγχάνουν καλύτερα αποτελέσματα από τα γραμμικά.
- Εντός του δίσκου τα τετραγωνικά και τα κυβικά στοιχεία είναι ελαφρώς χειρότερα αλλά η διαφορά κρίνεται αμελητέα, δίοτι η ακρίβεια είναι της τάξης του 10^{-4}.
- Υπάρχει τουλάχιστον μια τιμή για τα τετραγωνικά και τα κυβικά όπου εμφανίζεται μεγαλύτερη απόκλιση από ότι στα γραμμικά όμως η διαφορά κρίνεται αμελητέα, δίοτι η ακρίβεια είναι της τάξης του 10^{-4}.

Από τα παραπάνω προκύπτει ότι τα τετραγωνικά και τα κυβικά στοιχεία αντιμετωπίζουν το πρόβλημα της αστάθειας στην κυκλική συνοριακή περιοχή σαφώς καλύτερα με το τίμημα ότι εντός του δίσκου έχουμε αμελητέα μικρότερη ακρίβεια της τάξης του 10^{-4} και ότι υπάρχει ένα σημείο όπου η χείριστη προσέγγιση είναι χειρότερη αλλά και πάλι της τάξης του 10^{-4}, συνεπώς θεωρούμε ότι κρίνονται προτιμητέα στο σύνολο.

Ως τρίτο σημείο για την απόδοση του αλγορίθμου στο πρόβλημα Π1, σημειώνουμε ότι λαμβάνουμε μέτρια αποτελέσματα για χαμηλή ποιότητα εκλέπτυνσης όμως τα θεωρούμε αναμενόμενα όπως άλλωστε μπορούμε να διακρίνουμε και στα παρακάτω γραφήμα όπου είναι σαφές ότι μικρό πλήθος τριγώνων αδυνατεί να αντιμετωπίσει το κυκλικό σύνορο.

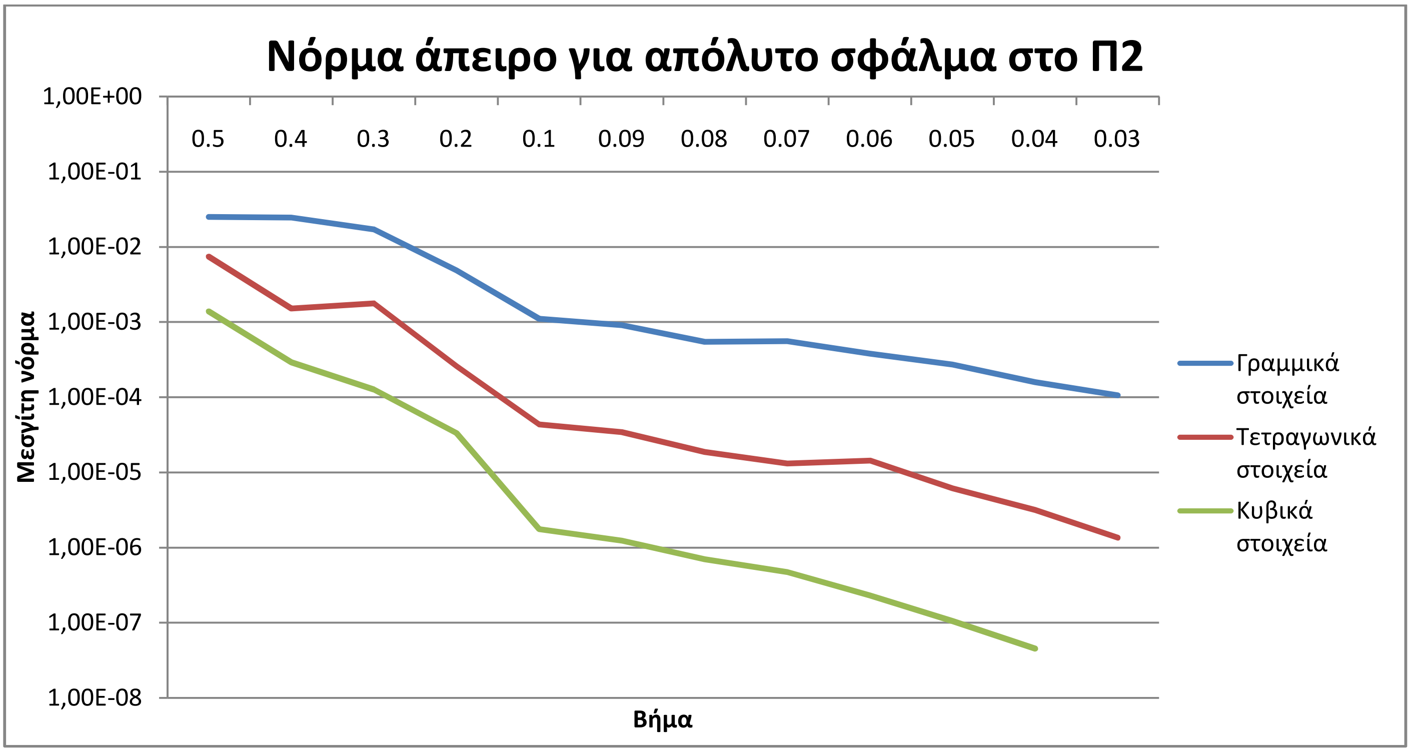


Γραμμικά στοιχεία για την τιμή της συνάρτησης, την εκτίμηση και το απόλυτο σφάλμα για h=0.5 στο Π1.

Συνοψίζοντας για την απόδοση του αλγορίθμου στο πρόβλημα εν γένει θεωρούμε ότι μπορεί να χαρακτηριστεί ιδιαίτερα ικανοποιητική, διότι τα μέτρια αποτελέσματα για χαμηλή ποιότητα εκλέπτυνσης είναι αναμενόμενα και καθώς το βήμα βελτιώνεται έχουμε ταχύτατα μετάβαση σε σφάλματα της τάξης του 10^{-4}, όπου εάν συνυπολογίσουμε την κλίση των ευθειών αναμένουμε βελτίωση για χαμηλότερες τιμές του h και περαιτέρω παρατηρούμε ειδικά για τα τετραγωνικά και τα κυβικά στοιχεία εξαιρετική ομοιομορφία και ευστοχία στην δύσκολη συνοριακή περιοχή.

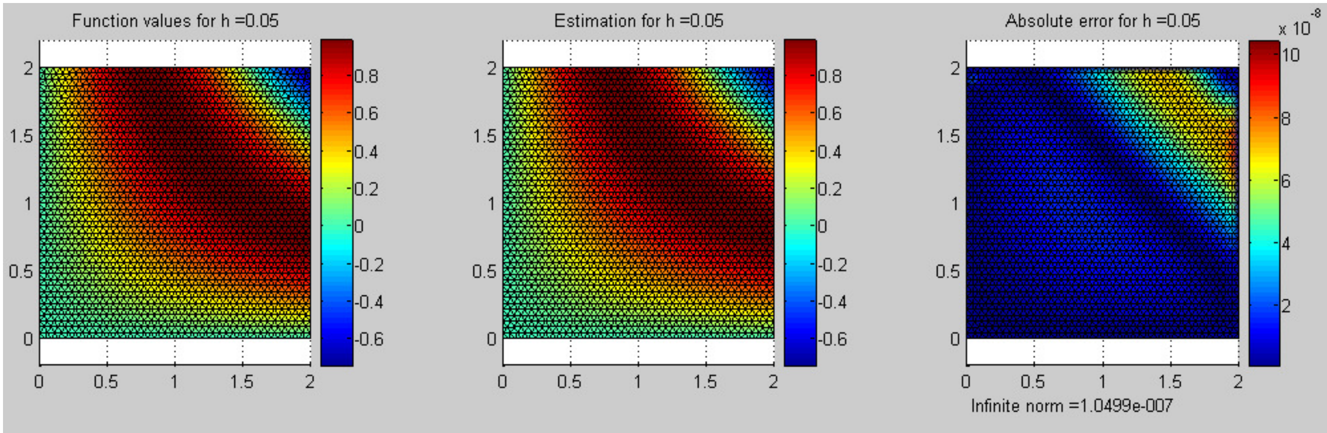
Π2	$-u_{xx} - u_{yy} = y^2 \sin(xy) + x^2 \sin(xy)$		
	Νόρμα άπειρο		
h	**Γραμμικά στοιχεία**	**Τετραγωνικά στοιχεία**	**Κυβικά στοιχεία**
0.5	2,50E-02	7,42E-03	1,38E-03
0.4	2,44E-02	1,52E-03	2,90E-04
0.3	1,70E-02	1,78E-03	1,26E-04
0.2	4,88E-03	2,59E-04	3,33E-05
0.1	1,11E-03	4,35E-05	1,75E-06
0.09	9,04E-04	3,42E-05	1,23E-06
0.08	5,45E-04	1,87E-05	6,96E-07
0.07	5,58E-04	1,31E-05	4,70E-07
0.06	3,80E-04	1,43E-05	2,29E-07
0.05	2,71E-04	6,09E-06	1,05E-07
0.04	1,59E-04	3,16E-06	4,50E-08
0.03	1,06E-04	1,35E-06	Πέρας αλγορίθμου

Πίνακας 5.8. Τιμές της νόρμα άπειρο στο Π2, για διαφορετικούς τύπους στοιχείων με βήμα έως και 0.03 .

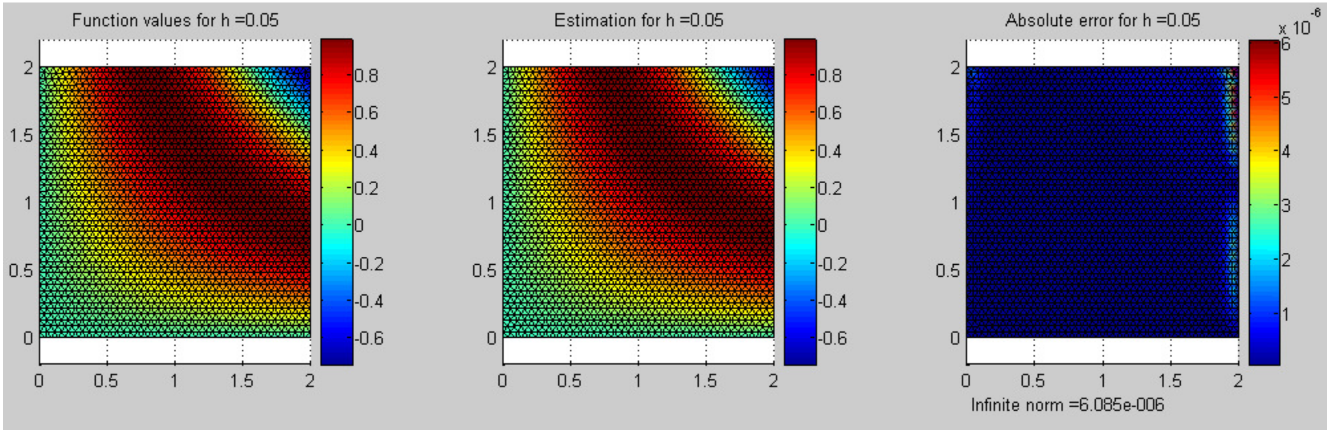


Μεγίστη νόρμα σε λογαριθμική κλίμακα για απόλυτο σφάλμα, με τρείς τύπους και διάφορες τιμές εκλέπτυνσης στο Π2.

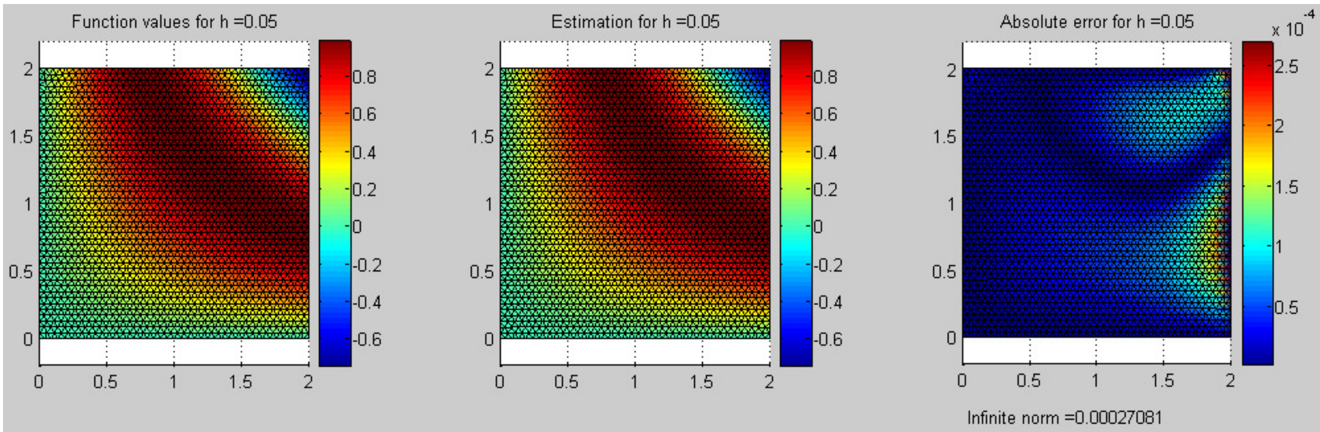
Στην περίπτωση του προβλήματος Π2, όπου πλέον δεν έχουμε κυκλικό σύνορο, παρατηρούμε ότι η απόδοση του αλγορίθμου είναι εξαιρετική με αποτελέσματα της τάξης του 10^{-8} και η βελτίωση στην προσέγγιση είναι ταχύτατη όπως φαίνεται και από το παραπάνω γράφημα. Περαιτέρω τα αποτελέσματα λαμβάνουν την αναμενόμενη μορφή όπου τα κυβικά στοιχεία είναι σαφώς καλύτερα από τα τετραγωνικά στοιχεία, τα οποία με την σειρά τους είναι σαφώς καλύτερα από τα γραμμικά με εξαιρετική ομοιομορφία στην απόδοση του αλγορίθμου (Σημειώνουμε ότι η κίτρινη περιοχή που εμφανίζεται για τα κυβικά στοιχεία είναι της τάξης του 10^{-8} ενώ για τα τετραγωνικά και τα γραμμικά στοιχεία η τάξη είναι 10^{-6} και 10^{-4} αντίστοιχα) .



Τιμή της συνάρτησης, της εκτίμησης και του απολύτου σφάλματος με κυβικά στοιχεία για h=0.05 στο Π2.



Τιμή της συνάρτησης, της εκτίμησης και του απολύτου σφάλματος με τετραγωνικά στοιχεία για h=0.05 στο Π2.



Τιμή της συνάρτησης, της εκτίμησης και του απολύτου σφάλματος με γραμμικά στοιχεία για h=0.05 στο Π2.

Συνοψίζοντας για τα αποτελέσματα μας, θεωρούμε ότι τα πεπερασμένα στοιχεία απέδωσαν υψηλής ακρίβειας προσεγγίσεις και για τις δυο μεθόδους σε πολλαπλές διαφορετικής μορφής συναρτήσεις και πεδία ορισμού με την μέθοδο Galekrin να επιβεβαιώνει τις βιβλιογραφικές αναφορές ως προτιμότερη ειδικά σε θέματα υπολογιστικού κόστους σε σφάλματα της τάξης 10^{-7} και παρακάτω. Ως μελλοντικές κατευθύνσεις θεωρούμε ότι η μίξη μορφής στοιχείων σε συνοριακές και μη περιοχές, π.χ. κυβικά στο σύνορο και γραμμικά εσωτερικά αλλά και το δυναμικό βήμα, π.χ. μικρές τιμές του h για σύνορο και πιο μεγάλες για δίσκο θα μας επιτρέψουν να έχουμε πολύ μεγάλη ακρίβεια με μικρό υπολογιστικό κόστος.

6. Παράρτημα – Κώδικας Matlab

Garerkin – 1d

```
function y = f(x)
% Ορισμός της υπο εκτίμηση συνάρτησης - αφαιρούμε τον σχολιασμό από την ζητούμενη
% y = 1;
% y = x;
% y = x^2;
% y = x^2+x+1;
% y = x^3;
% y = x^3+x^2+x+1;
% y = exp(x);
% y = exp(x)*x;
% y = exp(x)*x^2;
% y = exp(x)*(x^2+x+1);
% y = exp(x)*(x^3);
% y = exp(x)*(x^3+x^2+x+1);
% y = sin(x);
% y = sin(x)*exp(x);
return
```

```
function y = hat_one(x,x1,x2)
% Υπολογισμός της συνάρτησης hat με h=x2-x1 στο [x1,x2], με την συνάρτηση να
% λαμβάνει την τιμή ένα στο x2 και την τιμή μηδέν στο x1.
    y = (x-x1)/(x2-x1);
return
```

```
function y = hat_two(x,x1,x2)
% Υπολογισμός της συνάρτησης hat με h=x2-x1 στο [x1,x2], με την συνάρτηση να
% λαμβάνει την τιμή μηδέν στο x2 και την τιμή ένα στο x1.
    y = (x2-x)/(x2-x1);
return
```

```
function y = int_one(x1,x2)
% Υπολογισμός της συνεισφοράς στο load vector από την συνάρτηση hat_one, δηλαδή
% υπολογισμός του ορισμένου ολοκληρώματος με άκρα το x1,x2 και hat_one ως
% συνάρτηση προς ολοκλήρωση βάσει του κανόνα Simpson.
xmesos = (x1+x2)*0.5;
y   =   (x2-x1)*(f(x1)*hat_one(x1,x1,x2)   +   4*f(xmesos)*hat_one(xmesos,x1,x2)+
f(x2)*hat_one(x2,x1,x2) )/6;
return
```

```
function y = int_two(x1,x2)
% Υπολογισμός της συνεισφοράς στο load vector από την συνάρτηση hat_two, δηλαδή
% υπολογισμός του ορισμένου ολοκληρώματος με άκρα το x1,x2 και hat2 ως
% συνάρτηση προς ολοκλήρωση βάσει του κανόνα Simpson.
xm = (x1+x2)*0.5;
y      =      (x2-x1)*(f(x1)*hat_two(x1,x1,x2)      +      4*f(xmesos)*hat_two
(xmesos,x1,x2)+f(x2)*hat_two(x2,x1,x2) )/6;
return
```

```
function U = femcoef(x)
%--------- Υπολογισμός του βήματος h για το δοθέν πλέγμα -------------
M = length(x);
for i=1:M-1,
  h(i) = x(i+1)-x(i);
end
%--------------------------------------------------------------------

%--------- Αρχικοποίηση πινάκων A και F ------------------------------
A = sparse(M,M); F=zeros(M,1);
A(1,1) = 1; F(1)=0;
A(M,M) = 1; F(M)=0;
A(2,2) = 1/h(1); F(2) = int_one(x(1),x(2));
%--------------------------------------------------------------------

%---------Ορισμός πινάκων βάσει της μεθοδολογίας -------------------------
for i=2:M-2
  A(i,i)     = A(i,i) + 1/h(i);
  A(i,i+1)   = A(i,i+1) - 1/h(i);
  A(i+1,i)   = A(i+1,i) - 1/h(i);
  A(i+1,i+1) = A(i+1,i+1) + 1/h(i);
  F(i)       = F(i) + int_two(x(i),x(i+1));
  F(i+1)     = F(i+1) + int_one(x(i),x(i+1));
end
  A(M-1,M-1) = A(M-1,M-1) + 1/h(M-1);
  F(M-1)     = F(M-1) + int_two(x(M-1),x(M));
  U = A\F;
return
```

```
function u = real_values(x)
% Υπολογισμός των τιμών της συνάρτησης - αφαιρούμε τον σχολιασμό από την
% αντίστοιχη
% u = -(1/2)*x*(x-1);
% u = (1/6)*(x-x^3);
% u = (1/12)*(x-x^4);
% u = -(1/12)*x*(x^3+2*x^2+6*x-9);
% u = (1/20)*(x-x^5);
% u = -(1/60)*x*(3*x^4+5*x^3+10*x^2+30*x-48);
% u = exp(1)*x-x-exp(x)+1;
% u = -exp(x)*(x-2)+2*(x-1)-exp(1)*x;
% u = -exp(x)*(x^2-4*x+6)+3*exp(1)*x-6*x+6;
% u = -exp(x)*(x^2-3*x+5)+3*exp(1)*x-5*x+5;
% u = -exp(x)*(x^3-6*x^2+18*x-24)+24*(x-1)-11*exp(1)*x;
% u = -exp(x)*(x^3-5*x^2+15*x-19)+19*(x-1)-8*exp(1)*x;
% u = sin(x)-x*sin(1);
% u = (1/2)*(x-exp(1)*x*cos(1)+exp(x)*cos(x)-1);
```

```
function y = fem_values(x,U,xp)
M = length(x);
for i=1:M-1,
  if xp >=x(i) & xp <= x(i+1)
     y = hat_two(xp,x(i),x(i+1))*U(i) + hat_one(xp,x(i),x(i+1))*U(i+1);
     return
  end
end
```

Script για υπολογισμό αριθμητικών τιμών του προβλήματος testfem.m

```
% Καθαρισμός οθόνης και workspace
clc; clear all; close all;

% Παραγωγή πλέγματος
x = 0:0.005:1;
% Υπολογισμός των συντελεστών των πολυωνύμων στα κομβικά σημεία
U = femcoef(x);

% Εύρεση λύσεων από τα πεπερασμένα στοιχεία και σύγκριση τους με τις
% πραγματικές λύσεις (προσδιορισμός σφάλματος)
x2 = 0:0.005:1; k2 = length(x2);
for i=1:k2,
  u_real(i) = real_values(x2(i));
  u_fem(i) = fem_values(x,U,x2(i));
end

% Υπολογισμός της μετρικής άπειρο
error = norm(u_fem-u_real,inf)

% Γραφικές παραστάσεις
figure(1); plot(x2,u_fem-u_real);
xlabel('x'); title('Error plot')
```

Galerkin – 2d - Γραμμικά στοιχεία

```
% Καθαρισμός οθόνης, workspace ορισμός ρυθμού πτώσης του βήματος και δημιουργία
% αρχείων
clc;
close all;
step=-0.1;
fid=fopen('galerkin_p1_big.txt','a');
fid_2=fopen('galerkin_p1_big_norm.txt','a');

% Έναρξη αναδρομής (Εάν θέλουμε άλλες τιμές βήματος αλλάζουμε τα άκρα)
for rep=0.5:step:0.1

    % Χρησιμοποιούμε την βιβλιοθήκη distmesh η οποία μας επιστρέφει τις κορυφές
    % των τριγώνων ως μια τριάδα που δηλώνει ποιες είναι οι κορυφές ,π.χ. 1,2,3
    % και ένα διάνυσμα στήλη στο οποίο κάθε γραμμή αντιστοιχή στις συντεταγμένες
    % της i κορυφής [21]
    fd=@(p) sqrt(sum(p.^2,2))-1;
    [p,t]=distmesh2d(fd,@huniform,rep,[-1,-1;1,1],[]);
    be=boundedges(p,t);
    b=unique(be);

    % Υπο εκτίμηση συνάρτηση
    f=vectorize(inline('4','x','y'));
    k=vectorize(inline('1','x','y'));
    u=vectorize(inline('1-x^2-y^2','x','y'));
    ux=vectorize(inline('-2*x','x','y'));
    uy=vectorize(inline('-2*y','x','y'));

    % Το p που παράγεται από την distmesh έχει τις συντενταγμένες της N κορυφής
    % άρα το μέγεθος του θα είναι το πλήθος των κορυφών
    N=size(p,1);
```

```
% Το t που παράγεται από την distmesh έχει τις τριάδες των κορυφές οι οποίες
% απαρτίζουν κάθε τρίγωνο άρα το μέγεθος του θα είναι το πλήθος των τριγώνων
T=size(t,1);

% Αρχικοποίηση του stiffness matrix και του load vector
K=sparse(N,N); % zero matrix in sparse format: zeros(N) would be "dense"
F=zeros(N,1); % load vector F to hold integrals of phi's times load f(x,y)

% Έναρξη αναδρομής όπου ολοκληρώνουμε για όλα τα τρίγωνα
for i=1:T
  nodes=t(i,:);               % Οι κορυφές του i τριγώνου
  Pe=[ones(3,1),p(nodes,:)]; % 3 x 3 πίνακα για εύρεση εμβαδού
  Area=abs(det(Pe))/2;       % Εμβαδόν τριγώνου

  % Ο αντίστροφος του πίνακα Pe θα είναι η λύση του συστήματος όπως αναλύουμε
  % και στην μεθοδολογία – τροποποίση ψευδοκώδικα [5]
  C=inv(Pe);
  grad=C(2:3,:);                          % Οι τιμές b,c για το πολυώνυμο
  qpt = (1/3)*sum(p(nodes,:));            % Βαρύκεντρο
  I = feval(k,qpt(1),qpt(2));             % Υπολογισμός τιμής για το βαρύκεντρο
  g = feval(f,qpt(1),qpt(2));

  % Το τοπικό load vector θα είναι η σταθμισμένη τιμή το εμβαδόν συνεπώς
  % εφόσον έχουμε τρια σημεία για τα γραμμικά στοιχεία πολ/ζουμε με 3 το
  % βαρύκεντρο και σε κάθε σημείο θα αναλογεί το εμβαδόν / τιμή
  Fe=Area/3*g;
  % Ο τοπικός stiffness matrix πίνακας θα έινια το εμβαδόν επί
  % Επειδή χρειαζόμαστε μορφή πινάκων I * grad’ * grad
  Ke=Area*I*grad'*grad;

  % Προσθέτουμε κάθε τοπικό πίνακα με τους έως τώρα αποθηκευθέντες για
  % να κατασκευάσουμε το καθολικό load vector και το καθολικό stiffness
  % matrix
  K(nodes,nodes)=K(nodes,nodes)+Ke;    % Προσθήκη του τοπικού stiffness matrix
  F(nodes)=F(nodes)+Fe;                % Προσθήκη του τοπικού load vector
end

% Μηδενισμός στοιχείων για τις αντίστοιχες συνοριακές συνθήκες
K(b,:)=0; K(:,b)=0; F(b)=0; % put zeros in boundary rows/columns of K and F
K(b,b)=speye(length(b),length(b)); % put I into boundary submatrix of K

% Ο τελικός stiffness matrix και το τελικό load vector
Kb=K; Fb=F;

% Η λύση του συστήματος μας δίνει την εκτίμηση μας για τις κορυφές
U=Kb\Fb;

% Υπολογισμός των τιμών της συνάρτησης στις κορυφές – τροποποίηση αλγορίθμου
real_values = feval(u,p(:,1),p(:,2));
fem_values = U;

% Υπολογισμός του απολύτου σφάλματος
err=abs(U - real_values);

% Σε κάθε πείραμα παράγονται τρια υπογραφήματα
subplot(2,3,1);

% Σε κάθε πείραμα παράγονται τρια υπογραφήματα
trisurf(t,p(:,1),p(:,2),0*p(:,1),real_values,'edgecolor',
'k','facecolor','interp');
view(2),axis([-1 1 -1 1]),axis equal,colorbar   % Προβολή σε 2d
title(strcat(' Function values for h = ',num2str(rep)));  % Τίτλος γραφήματος
```

```
 subplot(2,3,2);  % Επόμενο υπογράφημα
    trisurf(t,p(:,1),p(:,2),0*p(:,1),fem_values,'edgecolor',
    'k','facecolor','interp');
    view(2),axis([-1 1 -1 1]),axis equal,colorbar     % Προβολή σε 2d
    title(strcat(' Estimation for h = ',num2str(rep)));     % Τίτλος γραφήματος

    subplot(2,3,3);     % Επόμενο υπογράφημα

    trisurf(t,p(:,1),p(:,2),0*p(:,1),err,'edgecolor','k','facecolor','interp');
    view(2),axis([-1 1 -1 1]),axis equal,colorbar     % Προβολή σε 2d
    title(strcat(' Absolute error for h = ',num2str(rep))); % Τίτλος γραφήματος

    % Ενσωμάτωση της μετρικής στο τρίτο υπογράφημα
    text(-1.0,-1.5,strcat(' Infinite norm =',num2str(norm(err,inf))));

    figure; % Παράγουμε καινούριο γράφημα για την επόμενη επανάληψη

    norm(err,inf) % Εκτίμηση της νόρμας άπειρο

    % Αποθήκευση αποτελεσμάτων
     fprintf(fid,'\n--------------------------------------------------------\n');
     fprintf(fid,'%4s %3.2f\n','h = ',rep);
     fprintf(fid,'--------------------------------------------------------\n');
     fprintf(fid,'%-11s %-11s %-11s \n','Πραγματική',' Εκτίμηση','  Απόλυτο
     Σφάλμα');
     for j = 1 : length(real_values)
         fprintf(fid,'%-11.10f ',real_values(j));
         fprintf(fid,'%-11.10f ',fem_values(j));
         fprintf(fid,'%-11.10f ',err(j));
         fprintf(fid,'\n');
     end
     fprintf(fid,'\n%-15s %-11.10f','Νόρμα άπειρο απολύτων σφαλμάτων =
     ',norm(err,inf));
     fprintf(fid,'\n--------------------------------------------------------\n');

     fprintf(fid_2,'\n%-11.10f',norm(err,inf));
end

% Κλείσιμο αρχείων για ορθή αποθήκευση
fclose(fid);
fclose(fid_2);
```

Galerkin – 2d - Τετραγωνικά στοιχεία - Τροποποίηση κώδικα

```
% Στα τετραγωνικά στοιχεία πέρα των τριών κορυφών έχουμε ακόμα τρια σημεία
% συνεπώς χρειάζεται να προσθέσουμε ενα στοιχείο σε κάθε μέσο κάθε πλευράς (πρίν
% από την γραμμή 182 στον κώδικα)
N=size(p,1);T=size(t,1);
M=sparse(N,N);
points = N+1;

% Έναρξη αναδρομής για όλα τα τρίγωνα – τροποποίηση αλγορίθμου
for i=1:T
  nodes=t(i,:);
  if (M(nodes(1,1),nodes(1,2)) == 0) % Πρώτη πλευρά
     M(nodes(1,1),nodes(1,2)) = points;
     M(nodes(1,2),nodes(1,1)) = points;
     p(points,:) = mean(p([nodes(1,1) nodes(1,2)],:)); % Προσθέτουμε στο μέσο
     points = points + 1;
  end
```

```
if (M(nodes(1,2),nodes(1,3)) == 0) % Δεύτερη πλευρά
     M(nodes(1,2),nodes(1,3)) = points;
     M(nodes(1,3),nodes(1,2)) = points;
     p(points,:) = mean(p([nodes(1,2) nodes(1,3)],:)); % Προσθέτουμε στο μέσο
     points = points + 1;
  end
  if (M(nodes(1,1),nodes(1,3)) == 0) % Τρίτη πλευρά
     M(nodes(1,1),nodes(1,3)) = points;
     M(nodes(1,3),nodes(1,1)) = points;
     p(points,:) = mean(p([nodes(1,1) nodes(1,3)],:)); % Προσθέτουμε στο μέσο
     points = points + 1;
  end

% Ο πίνακας t περιέχει τις κορυφές του τριγώνου, συνεπώς πρέπει να προσθέσουμε
% τις τρεις καινούριες κορυφές στις θέσεις 4,5,6
  t(i,4) = M(nodes(1,1),nodes(1,2));
  t(i,5) = M(nodes(1,2),nodes(1,3));
  t(i,6) = M(nodes(1,1),nodes(1,3));

end

% Κατ' αναλογία για τα συνοριακά σημεία
for i=1:size(be)
  nodes = be(i,:);
  if (M(nodes(1,1),nodes(1,2)) ~= 0)
    b = [b; M(nodes(1,1),nodes(1,2))];
  end
end
```

Galerkin – 2d - Κυβικά στοιχεία - Τροποίηση κώδικα

```
% Στα κυβικά στοιχεία πέρα των τριών κορυφών έχουμε ακόμα έξι σημεία
% συνεπώς χρειάζεται να προσθέσουμε ενα στοιχείο σε κάθε μέσο κάθε πλευράς (πρίν
% από την γραμμή 182 στον κώδικα)
N=size(p,1);T=size(t,1);M=sparse(N,N);
M1=sparse(N,N);
M2=sparse(N,N);
points = N+1;

% Έναρξη αναδρομής για όλα τα τρίγωνα – τροποποίηση αλγορίθμου
for i=1:T
  nodes=t(i,:);
  p(points,:) = mean(p(nodes(1,1:3),:));
  t(i,4) = points;
  points  = points +1;
  if (M1(nodes(1,1),nodes(1,2)) == 0)     % Πρώτη πλευρά
     lnodes = linspace(p(nodes(1,1),:),p(nodes(1,2),:),4);
     M1(nodes(1,1),nodes(1,2)) = points;
     M1(nodes(1,2),nodes(1,1)) = points;
     p(points,:) = lnodes(:,2);
     points  = points+1;
     M2(nodes(1,1),nodes(1,2)) = points;
     M2(nodes(1,2),nodes(1,1)) = points;
     p(points,:) = lnodes(:,3);
     points  = points +1;
  end
```

```
if (M1(nodes(1,2),nodes(1,3)) == 0) % Δεύτερη πλευρά
     lnodes = linspace(p(nodes(1,2),:),p(nodes(1,3),:),4);
     M1(nodes(1,2),nodes(1,3)) = points;
     M1(nodes(1,3),nodes(1,2)) = points;
     p(points,:) = lnodes(:,2);
     points  = points+1;
     M2(nodes(1,2),nodes(1,3)) = points;
     M2(nodes(1,3),nodes(1,2)) = points;
     p(points,:) = lnodes(:,3);
     points  = points+1;
  end
  if (M1(nodes(1,1),nodes(1,3)) == 0)     % Τρίτη πλευρά
     lnodes = linspace(p(nodes(1,1),:),p(nodes(1,3),:),4);
     M1(nodes(1,1),nodes(1,3)) = points;
     M1(nodes(1,3),nodes(1,1)) = points;
     p(points,:) = lnodes(:,2);
     points  = points+1;
     M2(nodes(1,1),nodes(1,3)) = points;
     M2(nodes(1,3),nodes(1,1)) = points;
     p(points,:) = lnodes(:,3);
     points  = points+1;
  end

% Ο πίνακας t περιέχει τις κορυφές του τριγώνου, συνεπώς πρέπει να προσθέσουμε
% τις τρεις καινούριες κορυφές στις θέσεις 5,6,7,8,9 (Στην 4 έχουμε τον μετρητή)

  t(i,5) = M1(nodes(1,1),nodes(1,2));
  t(i,6) = M2(nodes(1,1),nodes(1,2));
  t(i,7) = M1(nodes(1,2),nodes(1,3));
  t(i,8) = M2(nodes(1,2),nodes(1,3));
  t(i,9) = M1(nodes(1,1),nodes(1,3));
  t(i,10) = M2(nodes(1,1),nodes(1,3));
end

% Κατ' αναλογία για τα συνοριακά σημεία
for i=1:size(be)
  nodes = be(i,:);
  if (M1(nodes(1,1),nodes(1,2)) ~= 0)
    b = [b; M1(nodes(1,1),nodes(1,2))];
    b = [b; M2(nodes(1,1),nodes(1,2))];
  end
end
```

Galerkin – Μεταβολή υπο εξέταση συνάρτησης

```
% Στην περίπτωση θέλουμε να εκτελέσουμε το πείραμα σε μια οποιαδήποτε άλλη
% συνάρτηση αντικαθιστούμε τις γραμμές 168 – 172, π.χ. για την δεύτερη συνάρτηση
% που εξετάζουμε στην διπλωματική

% f=vectorize(inline('y^2*sin(x*y) + x^2*sin(x*y)','x','y'));
% k=vectorize(inline('1','x','y'));
% u=vectorize(inline('sin(x*y)','x','y'));
% ux=vectorize(inline('y*cos(x*y)','x','y'));
% uy=vectorize(inline('x*cos(x*y)','x','y'));
```

Ritz – 1d

```
% Καθαρισμός οθόνης και workspace
clear;
clc;

% Μεταβλητές ελέγχου, αρχική τιμή για το πλήθος των σημειών N και δημιουργία αρχείου
flag=1;
kindex=1;
N=0;
fid=fopen('ritz.txt','a');

% Εκίννηση αναδρομής και ορισμός υπο εξέτασης συνάρτησης - για μεταβολή αλλάζουμε τις f,p και q.
while flag
f = '2*pi^2*sin(pi*x)';
q = 'pi^2';
p = '1';

% Αύξηση πλήθους σημείων και αρχικοποίηση πινάκων - τροποποίηση ψευδοκώδικα [4]
N=N+10;
X = zeros(1,N+2);H = zeros(1,N+1);
Q = zeros(6,N+1);A = zeros(1,N+1);
B = zeros(1,N+1);C = zeros(1,N+1);
a_temp = zeros(1,N+1);b_matrix = zeros(1,N+1);
zeta = zeros(1,N+1);
z = zeros(1,N+1);
X(1) = 0;
X(N+2) = 1;

% Ορίζουμε το βήμα είναι ανάλογο του πλήθους των σημείων
step = 1/(N+1);
for J = 1 : N
    X(J+1) = J*step;
    H(J) = step;
end;
H(N+1) = step;
N1 = N-1;

% Υπολογίζουμε τους πίνακες Q τους οποίους θα χρησιμοποιήσουμε για να
% υπολογίσουμε τους πίνακες a_temp και b_temp και την συνάρτηση βιβλιοθήκης
% simpson
 for J = 2 : N
     Q(1,J-1) = simpson (1,X(J),X(J+1),q,f,p)/((H(J))*(H(J)));
     Q(2,J-1) = simpson (2,X(J-1),X(J),q,f,p)/((H(J-1))*(H(J-1)));
     Q(3,J-1) = simpson (3,X(J),X(J+1),q,f,p)/((H(J))*(H(J)));
     Q(4,J-1) = simpson (4,X(J-1),X(J),q,f,p)/((H(J-1))*(H(J-1)));
     Q(5,J-1) = simpson (5,X(J-1),X(J),q,f,p)/  H(J-1) ;
     Q(6,J-1) = simpson (6,X(J),X(J+1),q,f,p)/  H(J)   ;
 end
     Q(2,N)   = simpson (2,X(N),X(N+1),q,f,p)/((H(N))*(H(N)));
     Q(3,N)   = simpson (3,X(N+1),X(N+2),q,f,p)/((H(N+1))*(H(N+1)));
     Q(4,N)   = simpson (4,X(N),X(N+1),q,f,p)/((H(N))*(H(N)));
     Q(4,N+1) = simpson (4,X(N+1),X(N+2),q,f,p)/((H(N+1))*(H(N+1)));
     Q(5,N)   = simpson (5,X(N),X(N+1),q,f,p)/H(N)
     Q(6,N)    = simpson (6,X(N+1),X(N+2),Q,F,P)/H(N+1);
```

```
% Υπολογίζουμε τους πίνακες A και B βάσει των a_temp και b_temp όπως αναλύονται
%  στην μεθοδολογία [4]
 for J = 2 : N1 + 1
    a_temp (J-1) = Q(4,J-1)+Q(4,J)+Q(2,J-1)+Q(3,J-1);
    b_temp (J-1) = Q(1,J-1)-Q(4,J);
    B(J-1) = Q(5,J-1)+Q(6,J-1);
 end

    a_temp (N) = Q(4,N)+Q(4,N+1)+Q(2,N)+Q(3,N);
    B(N) = Q(5,N)+Q(6,N);

    A(1) = a_temp(1);
    zeta (1) = b_temp (1)/ a_temp(1);
    z(1) = B(1)/A(1);

 for J = 2 : N1
    A(J) = a_temp (J)- b_temp(J-1)* zeta (J-1);
    zeta (J) = b_temp(J)/A(J);
    z(J) = (B(J)- b_temp (J-1)*Z(J-1))/A(J);
 end

    A(N) = a_temp(N) - b_temp (N-1) * zeta (N-1);
    z(N) = (B(N)- b_temp(N-1)*Z(N-1))/A(N);

 % Υπολογίζουμε τον πίνακα C ο οποίοε είναι η λύση του συστήματος αρά και η λύση
 % των πεπερασμένων στοιχείων
 C(N) = z(N);
 for J = 1 : N1
    J1 = N - J;
    C(J1) = z(J1)- zeta(J1)*C(J1+1);
 end

 % Υπολογισμός των τιμών της συνάρτησης και της εκτίμησης
 fem_values=C(1:length(C)-1);
 real_values=sin(pi*X);
 real_values=real_values(2:length(real_values)-1);

 % Υπολογισμός του απολύτου σφάλματος
 err=abs(fem_values-real_values);

 % Υπολογισμός του ποσοστιαίου σφάλματος
 for i=1:length(fem_values)
     err_perc(i)=err(i)/real_values(i);
 end
 err_perc=err_perc*100;

 % Έλεγχος εκπλήρωσης κριτηρίου ακρίβειας και ανάλογα τερματισμός ή όχι
 if norm(err,inf)<=0.00000000001
     flag=0;
 end

 % Εάν δεν τερματίσει ο αλγόριθμος προβάλουμε το πλήθος των σημείων και των
 % αντίστοιχων τιμών του σφάλματος για να έχουμε εικόνα της εξέλιξης τους
αλγορίθμου
 disp('Πλήθος τιμών = ');
 N
 disp('Νόρμα άπειρο απολύτων σφαλμάτων = ');
 norm(err,inf)
 disp('Νόρμα άπειρο σχετικων σφαλμάτων = ');
 norm(err_perc,inf)
```

```
% Αποθηκεύουμε στο Ritz.txt τα δεδομένα με την σειρά που φαίνεται στην 4η fprintf
 fprintf(fid,'\n------------------------------------------------------------\n');
 fprintf(fid,'%4s %3d\n','N = ',N);
 fprintf(fid,'------------------------------------------------------------\n');
 fprintf(fid,'%-12s            %-12s            %-12s            %-12s            %-
 11s\n','X','Πραγματική','Εκτίμηση','Απόλυτο Σφάλμα','Ποσοστιαίο Σφάλμα');
 for j = 1 : N
     fprintf(fid,'%11.10f ',X(j));
     fprintf(fid,'%11.10f ',real_values(j));
     fprintf(fid,'%11.10f ',fem_values(j));
     fprintf(fid,'%11.10f ',err(j));
     fprintf(fid,'  %20.19f',err_perc(j));
     fprintf(fid,'\n');
 end
 fprintf(fid,'\n%-15s    %-11.10f','Νόρμα    άπειρο    απολύτων    σφαλμάτων    =
 ',norm(err,inf));
 fprintf(fid,'\n%-15s    %-11.10f','Νόρμα    άπειρο    σχετικων    σφαλμάτων    =
 ',norm(err_perc,inf));
 fprintf(fid,'\n------------------------------------------------------------\n');

 % Στην μεταβλητή keep αποθηκεύουμε ξεχωριστά τις τιμές της νόρμας για να έχουμε
 % εύκολη επεξεργασία όταν ολοκληρωθεί ο αλγόριθμος (Για να μην χρειάζεται να
 % σαρώσουμε ολόκληρο το txt
 keep(kindex,1)=N;
 keep(kindex,2)=norm(err,inf);
 keep(kindex,3)=norm(err_perc,inf);
 kindex=kindex+1;

end

% Κλείσιμο αρχείου για ορθή αποθήκευση
fclose(fid);
```

7. Βιβλιογραφία - Διαδικτυακές αναφορές

1. Μαρία Χρ. Γουσίδου – Κουτίτα, ‘’ Αριθμητικές Μέθοδοι με εφαρμογές στη Θεωρία Ελέγχου ‘’ , 2006.
2. Μαρία Χρ. Γουσίδου – Κουτίτα, Πανεπιστημιακές Παραδόσεις ‘’ Αριθμητικές Μέθοδοι με εφαρμογές στην Επίλυση Κανονικών και Μερικών Διαφορικών Εξισώσεων ‘’ , 2009.
3. Μαρία Χρ. Γουσίδου – Κουτίτα, Πανεπιστημιακές Παραδόσεις ‘’ Υπολοστικά Μαθηματικά ΙΙ‘’ , 2003.
4. Faires & Burden, Numerical Methods 7th Edition, Thomson & Brooks/Cole (2005).
5. Gilbert Strang, Computational Science and Engineering, Wellesley – Cambridge Press (2007).
6. Daryl L. Logan, First Coursein the Finite Element Method, Fourth Edition, Thomson, 2007
7. www4.ncsu.edu/~zhilin/TEACHING/MA587/chap7.pdf
8. http://www.math.ucdavis.edu/~hunter/pdes/ch3.pdf
9. http://www2.icmc.usp.br/~andcarva/sobolew.pdf
10. buzzard.ups.edu/courses/2007spring/projects/ott-paper-revised.pdf
11. http://jones.math.unibas.ch/~beilina/CDE/Lectures/Lecture5.pdf
12. www.ann.jussieu.fr/~frey/cours/UdC/ma691/ma691_ch7.pdf
13. http://www.math.umn.edu/~olver/num_/lnz.pdf
14. www.math.iit.edu/~fass/478578_Chapter_12.pdf
15. http://www.maths.ox.ac.uk/system/files/coursematerial/2012/2631/28/LA-7-12.pdf
16. www.math.uh.edu/~rohop/spring_05/downloads/Chapter3.pdf
17. http://mathinsight.org/differentiability_multivariable_theorem
18. http://mathinsight.org/differentiable_function_discontinuous_partial_derivatives
19. http://math-cs.aut.ac.ir/~shamsi/Ebooks/ch%2004%20of%20Sadiku.pdf
20. http://wwwmayr.informatik.tu-muenchen.de/konferenzen/Jass04/courses/2/Papers/fem_introduction_(tex).pdf
21. http://persson.berkeley.edu/distmesh/persson-thesis-color.pdf
22. http://scholar.lib.vt.edu/theses/available/etd-08022005-145837/unrestricted/Chapter3FiniteElementMethod.pdf
23. www.ann.jussieu.fr/~frey/cours/UdC/ma691/ma691_ch3.pdf
24. http://people.sc.fsu.edu/~jburkardt/m_src/fem_50/fem_50.html