Numerics Exercise 3 - Hannah Franklin

Ex25_3 FEM Sheet (Poisson):

Poisson system:
$$(1)\begin{cases} -\nabla^2 u = f \quad \text{on } (x,y)\in[0,1]^2 \\ f(x,y) = 2\pi^2 \sin(\pi x)\cos(\pi y) \\ u(0,y) = u(1,y) = 0 \\ \frac{\partial u}{\partial y}\Big|_{y=0} = \frac{\partial u}{\partial y}\Big|_{y=1} = 0 \end{cases}$$

Claim: exact solution is $u_e(x,y) = \sin(\pi x)\cos(\pi y)$

$$-\nabla^2 u = -\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2}$$

$$\begin{aligned} -\nabla^2 u_e &= \pi^2 \sin(\pi x)\cos(\pi y) + \pi^2 \sin(\pi x)\cos(\pi y) \\ &= 2\pi^2 \sin(\pi x)\cos(\pi y) \\ &= f \quad \text{as required.} \end{aligned}$$

$$\left.\begin{aligned} &u_e(0,y) = \sin(0)\cos(\pi y) = 0 \\ &u_e(1,y) = \sin(\pi)\cos(\pi y) = 0 \\ &\frac{\partial u_e}{\partial y}\Big|_{y=0} = -\pi \sin(\pi x)\sin(0) = 0 \\ &\frac{\partial u_e}{\partial y}\Big|_{y=1} = -\pi \sin(\pi x)\sin(\pi) = 0 \end{aligned}\right\} \therefore \text{satisfies required BCs}$$

Hence u_e is exact solⁿ of Poisson system (1).

Step 1: Ritz - Galerkin principle for (1).

$\Omega := [0,1]^2$, $\Gamma_1 := \{x=0,1\}$, $\Gamma_2 := \{y=0,1\}$, $\partial\Omega := \Gamma_1 \cup \Gamma_2$

solⁿ to system (1) minimises the functional: $I[u] = \iint_\Omega \left(\frac{1}{2}|\nabla u|^2 - uf\right) d\Omega$

Riesz' method: $\delta I = \frac{dI}{d\varepsilon} = \lim_{\varepsilon\to 0}\left(\frac{I[u+\varepsilon(\delta u)] - I[u]}{\varepsilon}\right) = 0$ (variation $(\delta u)(x,y)$)

$$\begin{aligned} I[u+\varepsilon\delta u] - I[u] &= \iint_\Omega \left(\tfrac{1}{2}|\nabla(u+\varepsilon\delta u)|^2 - (u+\varepsilon\delta u)f\right) d\Omega - \iint_\Omega \left(\tfrac{1}{2}|\nabla u|^2 - uf\right) d\Omega \\ &= \iint_\Omega \left(\tfrac{1}{2}|\nabla u + \varepsilon\nabla(\delta u)|^2 - \varepsilon(\delta u)f - \tfrac{1}{2}|\nabla u|^2\right) d\Omega \\ &= \iint_\Omega \left(\tfrac{1}{2}|\nabla u|^2 + \varepsilon\nabla u\cdot\nabla(\delta u) + \tfrac{\varepsilon^2}{2}|\nabla(\delta u)|^2 - \varepsilon(\delta u)f - \tfrac{1}{2}|\nabla u|^2\right) d\Omega \\ &= \iint_\Omega \left(\varepsilon\nabla u\cdot\nabla(\delta u) + \tfrac{\varepsilon^2}{2}|\nabla(\delta u)|^2 - \varepsilon(\delta u)f\right) d\Omega \end{aligned}$$

$$\begin{aligned} \Rightarrow \delta I &= \lim_{\varepsilon\to 0}\left(\iint_\Omega \left[\nabla u\cdot\nabla(\delta u) + \tfrac{\varepsilon}{2}|\nabla(\delta u)|^2 - (\delta u)f\right] d\Omega\right) \\ &= \iint_\Omega \left(\nabla u\cdot\nabla(\delta u) - \delta u f\right) d\Omega \end{aligned}$$

$\delta I = 0 \Rightarrow \iint_\Omega \left(\nabla u\cdot\nabla(\delta u) - \delta u f\right) d\Omega = 0$ → weak formulation

IBPs (integration by parts) $\Rightarrow -\iint_\Omega \left(\delta u \nabla^2 u + \delta u f\right) d\Omega + \int_{\partial\Omega} \left(\delta u \nabla u\cdot \underline{n}\right) dS = 0$

$$-\iint_\Omega \delta u(\nabla^2 u + f) d\Omega + \int_{\Gamma_1} \delta u(\nabla u\cdot\underline{n}) dS + \int_{\Gamma_2} \delta u(\nabla u\cdot\underline{n}) dS = 0$$

BCs $\Rightarrow \nabla u \cdot \underline{n} = 0$ on Γ_2

$\Rightarrow$ require variation $\delta u(x,y) = 0$ on Γ_1

$\Rightarrow -\iint_\Omega \delta u (\nabla^2 u + f) d\Omega = 0$ $\forall$ variations δu that satisfy $\delta u = 0$ on Γ_1

$\Rightarrow -\nabla^2 u - f = 0 \Rightarrow -\nabla^2 u = f$ on Ω w/ $\nabla u \cdot \underline{n} = 0$ on Γ_2

$\therefore$ yields system (1) w/ condition that variation $\delta u = 0$ on Γ_1.

($\delta u(0,y) = \delta u(1,y) = 0$)

weak formulation:

let $w(x,y)$ be test fnct.

multiply system (1) by w and integrate over Ω:

$$\iint_\Omega w(-\nabla^2 u - f)\, d\Omega = 0$$

IBPs $\Rightarrow \iint_\Omega (\nabla u \cdot \nabla w - wf) d\Omega + \int_{\partial\Omega} w (\nabla u \cdot \underline{n}) dS = 0$

$\nabla u \cdot \underline{n} = 0$ on Γ_2 $\Rightarrow \iint_\Omega (\nabla u \cdot \nabla w - wf)\, d\Omega + \int_{\Gamma_1} w(\nabla u \cdot \underline{n}) dS = 0$

$\Rightarrow$ test fnct $w(x,y) = 0$ on Γ_1

$\Rightarrow \iint_\Omega (\nabla u \cdot \nabla w - wf) d\Omega = 0$ is weak formulation.
w/ test fnct $w(x,y) = 0$ on Γ_1.

This is the same form as obtained through Riesz' method w/ the same condition on test fnct $w(x,y)$ as on variation $\delta u(x,y)$.

hence $\delta u = w$.

Step 2: algebraic or discrete Ritz-Galerkin principle.

$u(x,y) \approx u_h(x,y) = u_j \varphi_j(x,y)$ for $\varphi_j(x,y)$ compact global basis fncts

$w(x,y) \approx w_h(x,y) = w_j \varphi_j(x,y)$

take case where $w_j = \delta_{ij} \Rightarrow w_h(x,y) = \varphi_i(x,y)$ ($\varphi_i(0,y) = \varphi_i(1,y) = 0$)

this also guarantees BC on Γ_1 for u_h.

weak formulation: $\iint_\Omega (\nabla u \cdot \nabla w - wf)\, d\Omega = 0$

substitute in to weak formulation $\Rightarrow \iint_\Omega (\nabla u_h \cdot \nabla w_h - w_h f) d\Omega = 0$

(u_j constants) $\Rightarrow \iint_\Omega (u_j \nabla\varphi_j \cdot \nabla\varphi_i - \varphi_i f) d\Omega = 0$

$\Rightarrow u_j \iint_\Omega (\nabla\varphi_j \cdot \nabla\varphi_i) d\Omega = \iint_\Omega \varphi_i f\, d\Omega$

$\Rightarrow$ weak formulation becomes $A_{ij} u_j = b_i$

where $A_{ij} := \iint_\Omega (\nabla\varphi_j \cdot \nabla\varphi_i) d\Omega$, $b_i := \iint_\Omega \varphi_i f\, d\Omega$

where global basis fnct $\varphi_i(x,y)$ is s.t. $\varphi_i(0,y) = \varphi_i(1,y) = 0\ \forall i$.

now consider variation of $u_n(x,y)$ w/ variation fnct $\delta u_n(x,y)$

$$\delta I = \frac{dI}{d\varepsilon} = \lim_{\varepsilon\to 0}\left(\frac{I[u_n+\varepsilon\delta u_n]-I[u_n]}{\varepsilon}\right)$$

$$= \lim_{\varepsilon\to 0}\left(\frac{1}{\varepsilon}\left[\iint_\Omega\left(\tfrac{1}{2}|\nabla(u_n+\varepsilon\delta u_n)|^2-(u_n+\varepsilon\delta u_n)f\right)d\Omega - \iint_\Omega\left(\tfrac{1}{2}|\nabla u_n|^2-u_nf\right)d\Omega\right]\right)$$

$$= \lim_{\varepsilon\to 0}\left(\frac{1}{\varepsilon}\left[\iint_\Omega\left(\varepsilon\nabla u_n\cdot\nabla\delta u_n+\frac{\varepsilon^2}{2}|\nabla\delta u_n|^2-\varepsilon\delta u_n f\right)d\Omega\right]\right)$$

$$= \lim_{\varepsilon\to 0}\left(\iint_\Omega\left[\nabla u_n\cdot\nabla\delta u_n+\frac{\varepsilon}{2}|\nabla\delta u_n|^2-\delta u_n f\right]d\Omega\right)$$

$$= \iint_\Omega\left(\nabla u_n\cdot\nabla\delta u_n-\delta u_n f\right)d\Omega$$ this is weak formulation provided $\delta u_n = w_n$

require $\delta u_n(0,y)=\delta u_n(1,y)=0$ from step 1

step 1 $\Rightarrow \delta u_n = w_n = \varphi_i$

$\Rightarrow \iint_\Omega\left(u_j\nabla\varphi_j\cdot\nabla\varphi_i-\varphi_i f\right)d\Omega = \delta I = 0$

$\Rightarrow \iint_\Omega\left(\nabla\varphi_j\cdot\nabla\varphi_i\right)d\Omega\, u_j = \iint_\Omega \varphi_i f\, d\Omega$

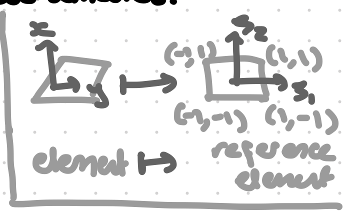
$\Rightarrow A_{ij}u_j = b_i$ w/ A_{ij}, b_i as defined previously.

hence the variation of u_n yields the algebraic eq. as required.

step 3: local coordinate system and reference coordinates.

quadrilateral elements: $N=4$ nodes $\underline{x}_{K,\alpha} = \underline{x}_\alpha$, $\alpha = 0,1,2,3$, element K

reference element $\underline{\xi} = (\xi_1,\xi_2) \in (-1,1)^2$



want to map $(x,y)\mapsto(\xi_1,\xi_2)$, $\underline{x}\mapsto\underline{\xi}$ | element $K \mapsto$ reference element $\hat{K}$

map coordinates of element K w/ shape fncts $\chi_\alpha(\underline{\xi})$ s.t.:

$$\underline{x} = \sum_{\alpha=0}^{3}\left[\underline{x}_{K,\alpha}\chi_\alpha(\underline{\xi})\right]$$, where $\underline{x}_{K,\alpha}$ local node coords of element K.

for quadrilaterals shape fncts:

$$\chi_0(\underline{\xi}) = \tfrac{1}{4}(1-\xi_1)(1-\xi_2)$$
$$\chi_1(\underline{\xi}) = \tfrac{1}{4}(1+\xi_1)(1-\xi_2)$$
$$\chi_2(\underline{\xi}) = \tfrac{1}{4}(1+\xi_1)(1+\xi_2)$$
$$\chi_3(\underline{\xi}) = \tfrac{1}{4}(1-\xi_1)(1+\xi_2)$$

these are also Galerkin basis fncts.

Jacobian of transformation $\underline{x} \mapsto \underline{\xi}$: $J = \begin{pmatrix} \frac{\partial x}{\partial \xi_1} & \frac{\partial x}{\partial \xi_2} \\ \frac{\partial y}{\partial \xi_1} & \frac{\partial y}{\partial \xi_2} \end{pmatrix}$

$$\underline{x} = \underline{x}_0 \chi_0 + \underline{x}_1 \chi_1 + \underline{x}_2 \chi_2 + \underline{x}_3 \chi_3$$

$$\Rightarrow J = \frac{1}{4} \begin{pmatrix} (1-\xi_2)(x_1 - x_0) + (1+\xi_2)(x_2 - x_3) & (1-\xi_1)(x_3 - x_0) + (1+\xi_1)(x_2 - x_1) \\ (1-\xi_2)(y_1 - y_0) + (1+\xi_2)(y_2 - y_3) & (1-\xi_1)(y_3 - y_0) + (1+\xi_1)(y_2 - y_1) \end{pmatrix}$$

from which $\det J$ and $(J^T)^{-1}$ can be found

$$\underline{\xi} = J \underline{x} \Rightarrow \underline{\nabla} \mapsto (J^T)^{-1} \begin{pmatrix} \frac{\partial}{\partial \xi_1} \\ \frac{\partial}{\partial \xi_2} \end{pmatrix}$$

$$d\underline{\xi} = \frac{1}{|\det J|} d\Omega \Rightarrow d\Omega = |\det J| d\underline{\xi}$$

considering $A_{ij} u_j = b_i$, $A_{ij} = \iint_\Omega (\nabla \varphi_j \cdot \nabla \varphi_i)\, d\Omega$, $b_i = \iint_\Omega \varphi_i f\, d\Omega$

χ_i are basis fncts $\Rightarrow$ $\hat{A}_{\alpha\beta} = \int_{\hat{K}} \left[(J^T)^{-1} \begin{pmatrix} \frac{\partial}{\partial \xi_1} \\ \frac{\partial}{\partial \xi_2} \end{pmatrix} \chi_\alpha \cdot (J^T)^{-1} \begin{pmatrix} \frac{\partial}{\partial \xi_1} \\ \frac{\partial}{\partial \xi_2} \end{pmatrix} \chi_\beta \right] |\det J| d\underline{\xi}$

$$\hat{b}_\alpha = \int_{\hat{K}} \left(f(\underline{\xi})\, \chi_\alpha\, |\det J| \right) d\underline{\xi}$$

for $\alpha, \beta = 0, 1, 2, 3$ on each reference element $\hat{K}$.

$\hat{A}_{\alpha,\beta}$, $\hat{b}_\alpha$ are elemental matrices and vectors.

global matrix A, global vector b

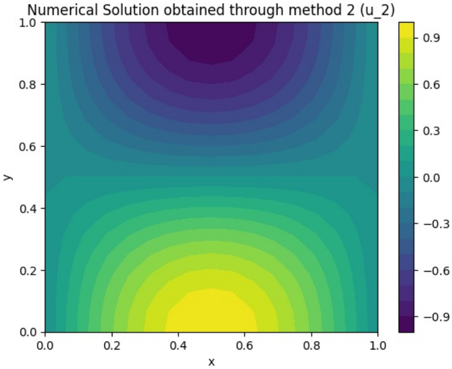
- set all elements of A, b to be 0. ($A_{ij} = b_i = 0 \ \forall i,j$)
- loop over all elements K : $K = 0, N_{el} - 1$ (N_{el} = no. elements)

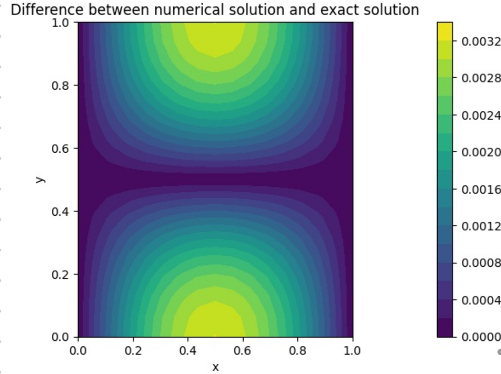
```
for α = 0, 3 :
    i = Index(K, α)
    for β = 0, 3 :
        j = Index(K, β)
        A_ij = A_ij + Â_αβ
    b_i = b_i + b̂_α
```

to define the global matrix using the elemental matrices.

4. Step 4: solve the system

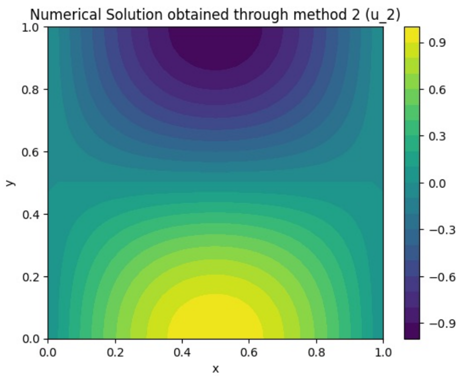
$\{h, p\} = \{\frac{1}{16}, 1\}$ contours of solution u_2 and $|u_2 - u_e|$:

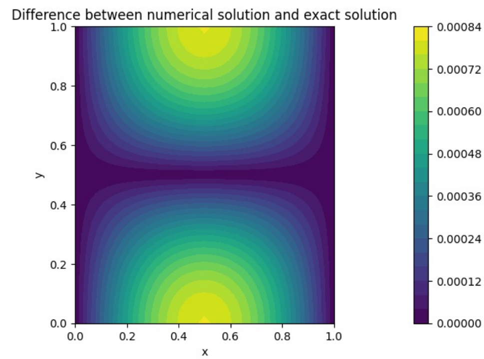




max error ÷4

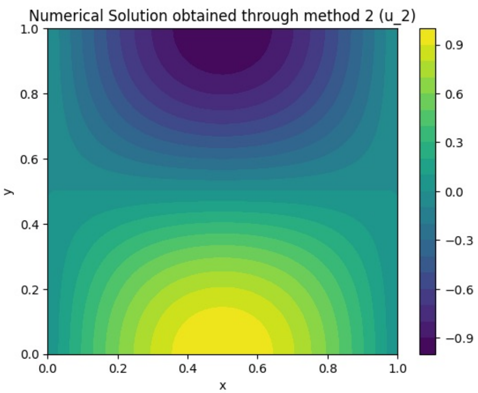
$\{h, p\} = \{\frac{1}{32}, 1\}$ contours:

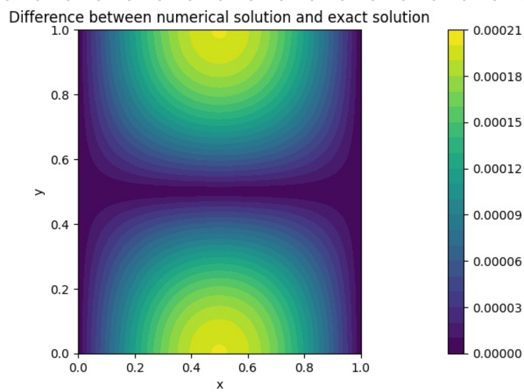




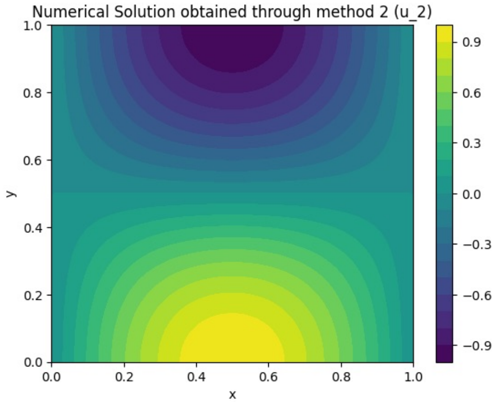
max error ÷4

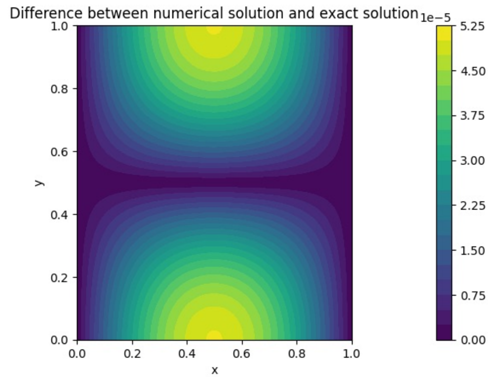
$\{h, p\} = \{\frac{1}{64}, 1\}$ contours:



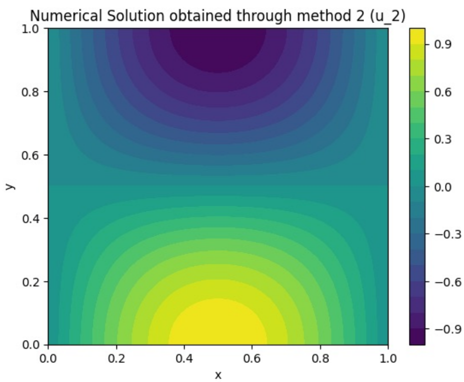


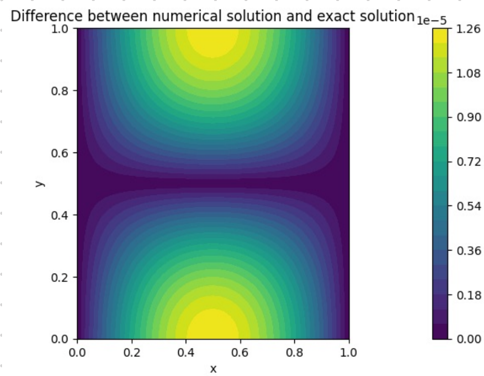
$\{h, p\} = \{\frac{1}{128}, 1\}$ contours:



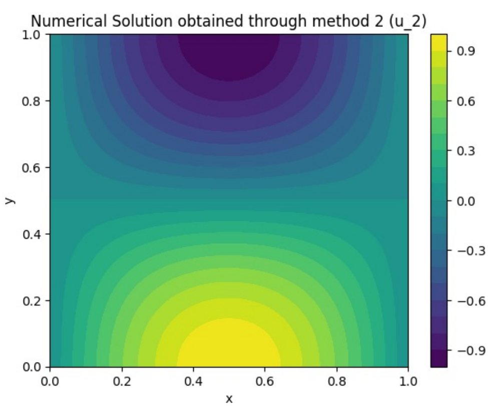


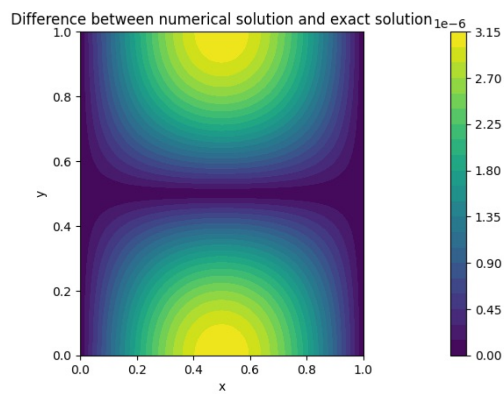
$\{h, p\} = \{\frac{1}{256}, 1\}$ contours:





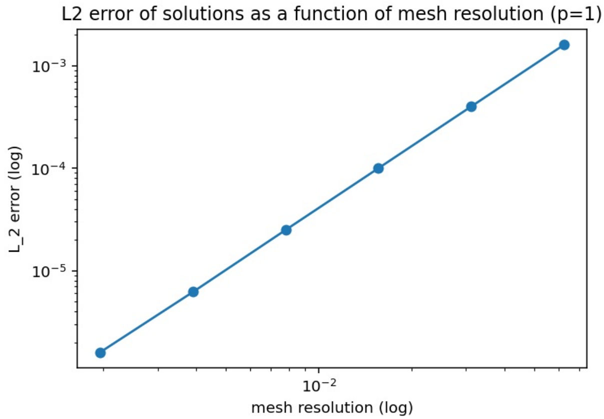
$\{h, p\} = \{\frac{1}{512}, 1\}$ contours:





L_2 error for different mesh resolutions (h values) for p=1.

h	L_2 error
$\frac{1}{16}$	1.6×10^{-3}
$\frac{1}{32}$	4.0×10^{-4}
$\frac{1}{64}$	1.0×10^{-4}
$\frac{1}{128}$	2.5×10^{-5}
$\frac{1}{256}$	6.2×10^{-6}
$\frac{1}{512}$	1.6×10^{-6}



Plotting the L_2 error of the solution as a function of the mesh resolution h for order p=1 allows us to find the dependence of the solution on h for p=1.

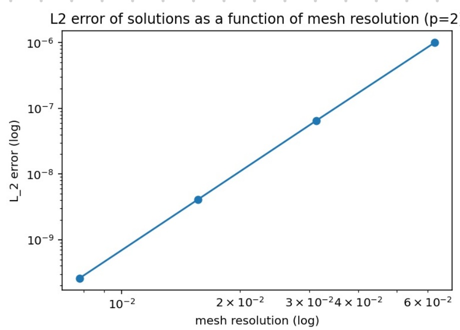
if $L_2 \propto h^\alpha$, taking the logscale of the plot allows us to calculate the value of α.

find that $\alpha = 1.996 \approx 2$

Thus $L_2 \propto h^2$ for p=1. => second order convergence in space for p=1.

Same analysis for p=2:

h	L_2 error
$\frac{1}{16}$	1.0×10^{-6}
$\frac{1}{32}$	6.5×10^{-8}
$\frac{1}{64}$	4.1×10^{-9}
$\frac{1}{128}$	2.6×10^{-10}



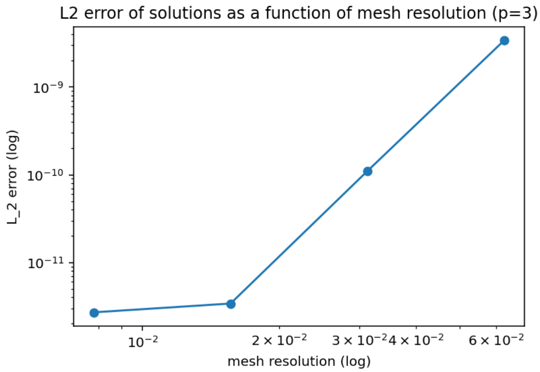
this time it is found that α: $3.97 \approx 4$

thus for $p=2$: $L_2 \propto h^4$ => 4th order convergence in space for $p=2$.

This suggests that $L_2 \propto h^{2p}$ ($p=1 \Rightarrow 2p=2$, $p=2 \Rightarrow 2p=4$)

For $p=3$:

h	L_2 error
$\frac{1}{16}$	3.4×10^{-9}
$\frac{1}{32}$	1.1×10^{-10}
$\frac{1}{64}$	3.4×10^{-12}
$\frac{1}{128}$	2.7×10^{-12}



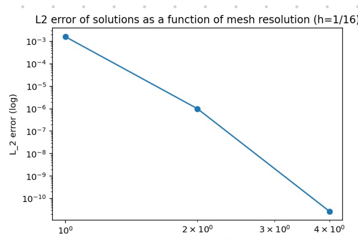
However this trend does not continue for $p=3$.
The linear portion of the graph has a gradient of ~5 which is closer to expected, but not the expected value of 6.

However at $p=3$ the errors are small enough that round-off error when generating the numerical solution becomes a concern.

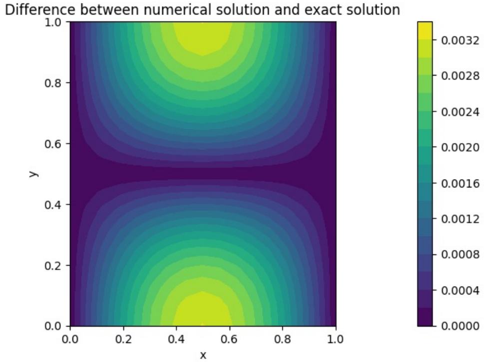
With this in mind, consider different values of p when $h = \frac{1}{16}$ to minimise the effect of round-off error.

$h = \frac{1}{16}$

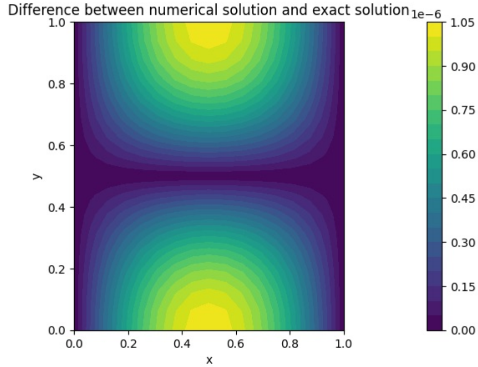
p	L_2 error
1	1.6×10^{-3}
2	1.0×10^{-6}
4	2.7×10^{-11}



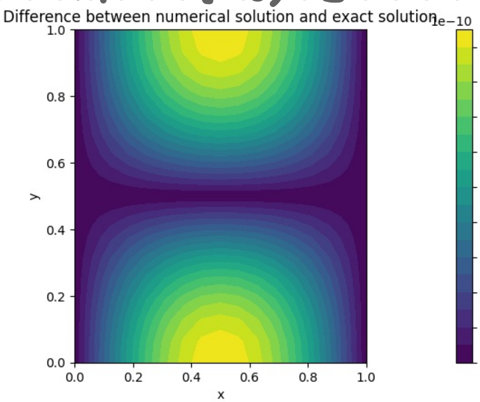
$\{h, p\} = \{\frac{1}{16}, 1\}$



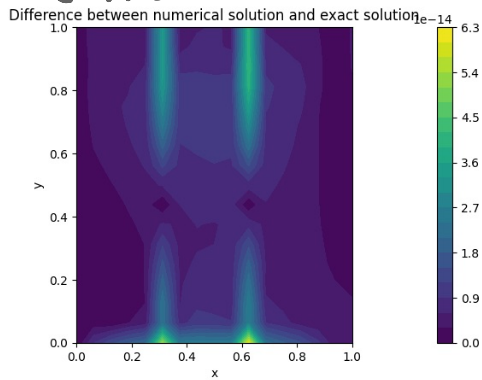
$\{h, p\} = \{\frac{1}{16}, 2\}$



$\{h, p\} = \{\frac{1}{16}, 3\}$



$\{h, p\} = \{\frac{1}{16}, 4\}$



(p=3 not included in analysis since consider doubling p)

if $L_2 \propto h^{2p}$ would expect that doubling p would square the L_2 error.

$p \to 2p \Rightarrow L_2 \propto h^{4p} = (h^{2p})^2$

$(1.6 \times 10^{-3})^2 = 2.56 \times 10^{-6} \sim 1.0 \times 10^{-6}$ if consider potential effects of round-off error.

$p = 1, 2$ $\quad h = \frac{1}{32}$: $\quad p = 1$: $L_2 = 4.0 \times 10^{-4}$

$(4.0 \times 10^{-4})^2 = 1.6 \times 10^{-7} \sim 6.5 \times 10^{-8}$

This is approximately the case, thus it can be conclude that the error scales w/ h^{2p}

$L_2 \sim h^2 p$ $\Rightarrow$ $\{h, p\}$ combinations w/ similar $h^2 p$ values are equivalent.

$h = 2^{-4}, 2^{-5}, 2^{-6}, 2^{-7}, 2^{-8}, 2^{-9}$ $\left(\frac{1}{16}, \frac{1}{32}, \frac{1}{64}, \frac{1}{128}, \frac{1}{256}, \frac{1}{512}\right)$

$p = 1, 2, 3, 4$

$\underbrace{(2^{-4})^2 = 2^{-8} \Rightarrow \left\{\frac{1}{16}, 2\right\} \sim \left\{\frac{1}{256}, 1\right\}}_{\text{example of equivalent } \{h,p\} \text{ pairs w/ same value of } h^2p}$

$L_2: 1.0\times10^{-6} \sim 6.3\times10^{-6}$

same order $\Rightarrow$ roughly equivalent

5. **Explain how first four steps are implemented in Firedrake.**

Uploaded code file is fully commented to identify where these steps occur in the code.

6. **Change and implement BCs for different $f(x,y)$ and exact solution $u(x,y)$.**

$-\nabla^2 u = f(x,y)$

consider $u_{exact} = \sin(\frac{\pi}{2}x)\cos(\pi y)$

then $u(0,y) = 0$

$u(1,y) = \sin(\frac{\pi}{2})\cos(\pi y)$
$= \cos(\pi y)$

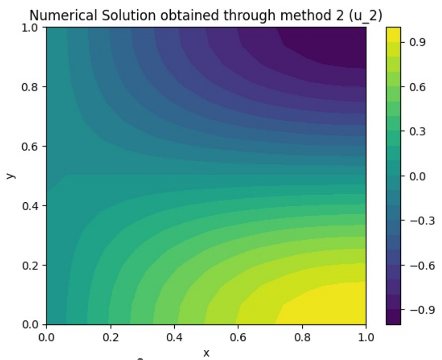
$$-\nabla^2 u = -\left(-\frac{\pi^2}{4}\sin(\tfrac{\pi}{2}x)\cos(\pi y) - \pi^2 \sin(\tfrac{\pi}{2}x)\cos(\pi y)\right)$$
$$= \frac{5\pi^2}{4}\sin(\tfrac{\pi}{2}x)\cos(\pi y)$$
$$= f(x,y)$$

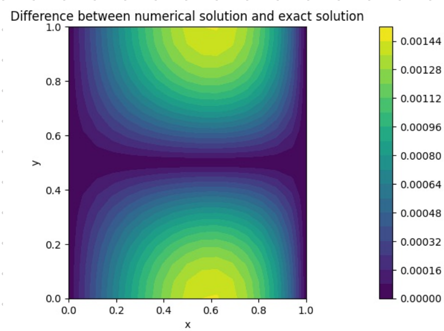
note: neumann BCs imposed, same as before.

Steps taken to change the code:

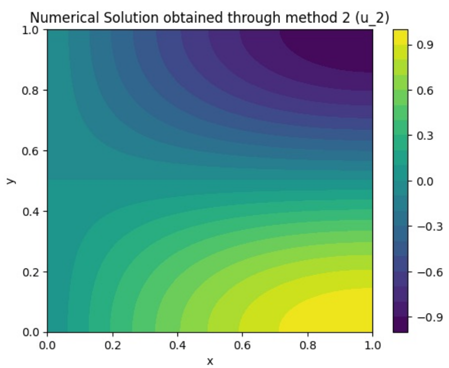
1. change function f to $\frac{5\pi^2}{4}\sin(\frac{\pi}{2}x)\cos(\pi y)$
2. change BCs bc_x_0, bc_x_1.
   bc_x_0 stays the same for this chosen function
   bc_x_1 no longer a constant, so function $\cos(\pi y)$ must be implemented instead.
3. change exact solution to $\sin(\frac{\pi}{2}x)\cos(\pi y)$
4. run code over chosen $\{h,p\}$ to veify it works.

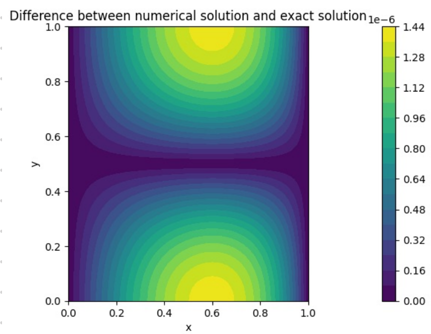
$\{h,p\} = \{\frac{1}{16}, 1\}$



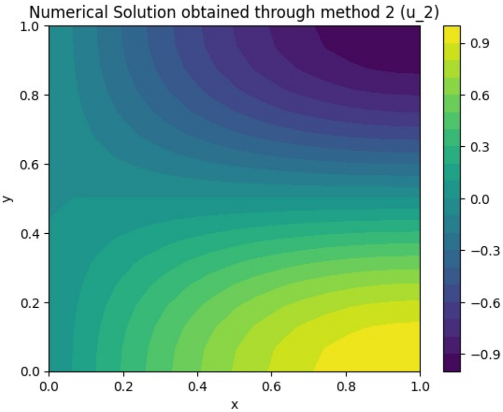


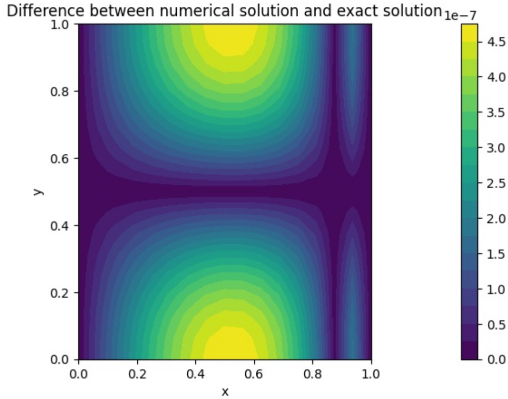
$\{h,p\} = \{\frac{1}{512}, 1\}$



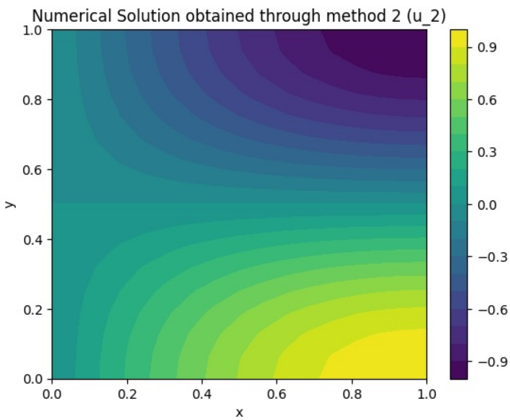


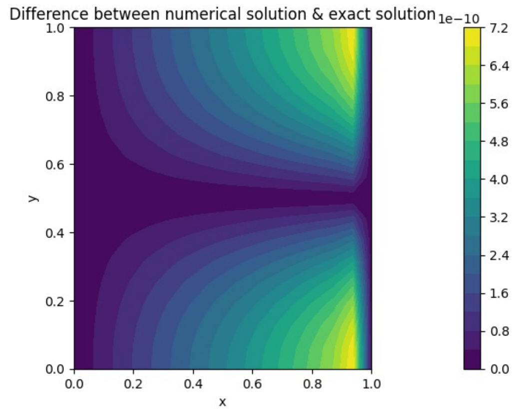
$\{h, p\} = \{\frac{1}{16}, 2\}$



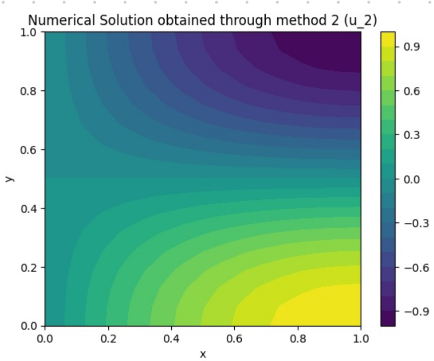


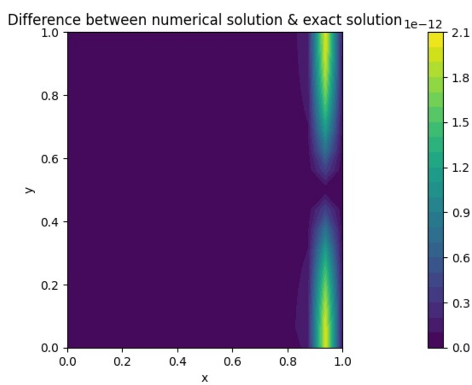
$\{h, p\} = \{\frac{1}{16}, 3\}$





$\{h, p\} = \{\frac{1}{16}, 4\}$





The error $|u_2 - u_{exact}|$ decreases w/ decreasing h
& increasing p

The error converges more quickly w/ changing p, as for the previously used f.

The contours for $|u_2 - u_{exact}|$ for the chosen $\{h, p\}$ pairs suggests that the new function $f(x,y)$ and BCs have been successfully implemented.