Math 5484, Finite Element Methods

Tao Lin
Department of Mathematics
Virginia Tech

These Lecture Slides are based on

Book 1: *Numerical Approximation of Partial Differential Equations*
by Soren Bartels

Book 2: *The Finite Element Methods*
Theory, Implementation and Applications
by Mats G. Larson and Fredrik Bengzon

Some prototypical differential equation problems

An 1D stationary heat equation: Consider a metal rod of length 1 with a diffusion coefficient k. Assume it is heated by a heat source f for a time period long enough so the the heat transfer within the rod has reached the steady state. What is the temperature distribution in this rod?

Let $T(x)$ be the temperature at the location $x \in I = [0, 1]$, $q(x)$ be the heat flux in the positive x direction, and $A(x)$ be the cross-section area of the rod at position x.

The first law of thermodynamics/balance of energy:

$$A(x+h)q(x+h) - A(x)q(x) = \int_x^{x+h} f(s)ds, \quad \forall x \in (0,1) \text{ and for a small } h$$

This leads to

$(A(x)q(x))' = f(x), \quad x \in (0,1)$ (a law involving derivatives)

Fourier's law: $q = -kT'$ (another law involving derivatives)

Heat equation: $-(AkT')' = f$ for $x \in (0,1)$

Physical laws involving derivatives lead to differential equations about a physical quantity $T(x)$ =???

This heat equation obviously does not have a unique solution for a given heat source f. Two more extra conditions are needed to uniquely determined a solution with the heat equation. These conditions are often given at end points of the rod or the interval.

Dirichlet boundary condition prescribes the value of temperature at the boundary, such as $T(L) = 0$. This is often called the essential or strong boundary condition.

Neumann boundary condition prescribes the derivative of the temperature at the boundary, such as $T'(0) = 0$, which means the rod is thermally isolated at left end because the heat flux $q(0) = -kT'(0) = 0$.

Robin boundary condition is a combination of Dirichlet and Neumann boundary condition: $AkT' = k(T - T_\infty) + Q_\infty$. where Q_∞ represents a heat flux entering the rod and $k(T - T_\infty)$ models the convective heat transfer where T_∞ is the ambient bulk temperature.

A boundary value problem (BVP) for the heat equation: find $T(x)$ such that

$$\begin{aligned}
&\text{the differential equation :} && -(AkT'(x))' = f(x),\ x \in (0,1)\\
&\text{the boundary condition at } x = 0 : && A(0)k(0)T'(0) = k_0(T(0) - g_0)\\
&\text{the boundary condition at } x = 1 : && -A(1)k(1)T'(1) = k_1(T(1) - g_1)
\end{aligned}$$

We will use the following set of functions defined on a set $D \subset \mathbb{R}^d,\ d = 1, 2, 3$:

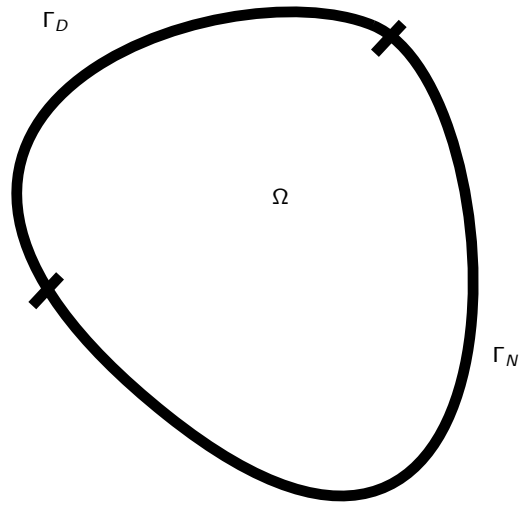
$$C^p(D) = \{u \mid u \text{ is a } p\text{-times continuously differentiable function}\}$$

Recall the BVP: find $T(x)$ such that

$$\begin{aligned}
&\text{the differential equation :} && -(AkT'(x))' = f(x),\ x \in (0,1)\\
&\text{the boundary condition at } x = 0 : && A(0)k(0)T'(0) = k_0(T(0) - g_0)\\
&\text{the boundary condition at } x = 1 : && -A(1)k(1)T'(1) = k_1(T(1) - g_1)
\end{aligned}$$

Definition 2.1 (BK1) $T(x)$ is a classical solution of the BVP above provided that $T(x)$ solves the differential equation in the BVP, satisfies the boundary condition in the BVP, and $T \in C^1([0,1]) \cap C^2((0,1))$.

A model stationary heat transfer problem in higher dimension: Consider an object occupying the domain $\Omega \subset \mathbb{R}^d, d = 2, 3$ with a diffusion coefficient k. Assume that this object has reached the thermal steady state with a source f.



Then the temperature distribution in Ω can be modeled by the following BVP: find $T(X)$ such that

$$\begin{aligned} \text{the PDE}: &\quad -\nabla \cdot (k \nabla T(X)) = f(X),\ X = (x, y) \in \Omega \\ \text{the boundary condition on } \Gamma_D: &\quad T|_{\Gamma_D} = g_D \\ \text{the boundary condition on } \Gamma_N: &\quad -\mathbf{n} \cdot \nabla T|_{\Gamma_N} = g_N \end{aligned}$$

where $\Gamma = \partial\Omega$ is the boundary of Ω and $\Gamma_D \cup \Gamma_N = \Gamma$.

A nonlinear heat transfer problem: find $T(X)$ such that

$$\begin{aligned} -\nabla \cdot (k(T)\nabla T) &= f(T),\ X \in \Omega \\ T(X) &= g(X),\ X \in \partial\Omega \end{aligned}$$

Of course, other types of boundary conditions can be considered.

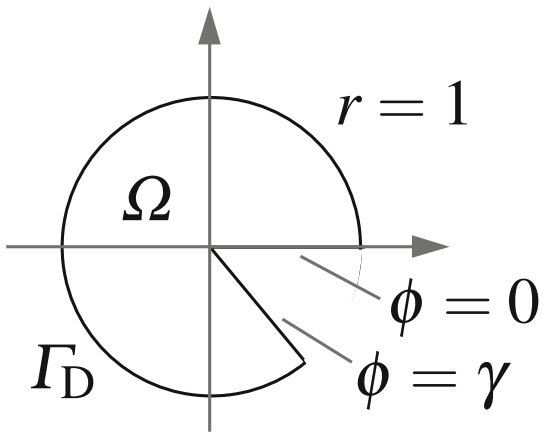
Consider BVP: find $T(X)$ such that

$$\text{the PDE :} \qquad -\triangle T = -\nabla \cdot (\nabla T(X)) = 0, \ X = (x, y) \in \Omega$$

$$\text{the boundary condition :} \qquad T|_{\partial\Omega} = g_D$$

and recall:

Definition 2.1 (BK1) $T(x)$ is a classical solution of the BVP above provided that $T(x)$ solves the differential equation in the BVP, satisfies the boundary condition in the BVP, and $T \in C^1([0,1]) \cap C^2((0,1))$.



For $\gamma \in (0, 2\pi)$, let

$$\Omega = \{(r\cos(\phi), r\sin(\phi)), \quad 0 < r < 1, \quad 0 < \phi < \gamma\}$$

and $\Gamma_D = \partial\Omega$, $f(X) = 0$, and

$$g_D(X) = g_D(r, \phi) = \begin{cases} 0, & \text{for } \phi = 0 \text{ or } \gamma \\ \sin(\phi\pi/\gamma), & \text{for } r = 1 \end{cases}$$

Proposition 2.1 (BK1): Then $u(X) = u(r, \phi) = r^{\pi/\gamma}\sin(\phi\pi/\gamma)$ is a classical solution of the BVP if and only if $\gamma \in (0, \pi]$.

Proof: We can easily verify that $u|_{\partial\Omega} = g_D$. Then

$$\begin{aligned}
u(X) &= u(r,\phi) = r^{\pi/\gamma}\sin(\phi\pi/\gamma) \\
\partial_r u(r,\phi) = u_r(r,\phi) &= \frac{\pi r^{-1+\pi/\gamma}}{\gamma}\sin(\phi\pi/\gamma) \\
\partial_r^2 u(r,\phi) = u_{rr}(r,\phi) &= \frac{\pi(\pi-\gamma)r^{-2+\pi/\gamma}}{\gamma^2}\sin(\phi\pi/\gamma) \\
\partial_\phi u(r,\phi) &= \frac{\pi r^{\pi/\gamma}}{\gamma}\cos(\phi\pi/\gamma) \\
\partial_\phi^2 u(r,\phi) &= -\frac{\pi^2 r^{\pi/\gamma}}{\gamma^2}\sin(\phi\pi/\gamma) \\
\text{Then } \triangle u &= \partial_x^2 u + \partial_y^2 u = \partial_r^2 u + \frac{1}{r}\partial_r u + \frac{1}{r^2}\partial_\phi^2 u = 0
\end{aligned}$$

which means u is a solution to the PDE: $-\triangle u = 0$. Also, we can show that $u \in C^\infty(\Omega) \subset C^2(\Omega)$. However,

$$\nabla u = \begin{bmatrix}\partial_x u \\ \partial_y u\end{bmatrix} = \begin{bmatrix}\partial_r u \\ \frac{1}{r}\partial_\phi u\end{bmatrix} = \left(\frac{\pi}{\gamma}\right) r^{\pi/\gamma-1}\begin{bmatrix}\sin(\phi\pi/\gamma) \\ \cos(\phi\pi/\gamma)\end{bmatrix}$$

Hence, ∇u is bounded on $\overline{\Omega}$ if and only if $\gamma \in (0,\pi]$, i.e., $u \in C^1(\overline{\Omega})$ if and only if $\gamma \in (0,\pi]$.

Recall the BVP: find $T(X)$ such that

$$\text{the PDE :} \qquad -\triangle T = -\nabla \cdot (\nabla T(X)) = 0,\ X = (x, y) \in \Omega$$
$$\text{the boundary condition :} \qquad T|_{\partial\Omega} = g_D$$

where, for $\gamma \in (0, 2\pi)$,

$$\begin{aligned} \Omega &= \{(r\cos(\phi), r\sin(\phi)),\ \ 0 < r < 1,\ \ 0 < \phi < \gamma\} \\ g_D(X) &= g_D(r, \phi) = \begin{cases} 0, & \text{for } \phi = 0 \text{ or } \gamma \\ \sin(\phi\pi/\gamma), & \text{for } r = 1 \end{cases} \end{aligned}$$

In summary:

(1) When $0 < \gamma \leq \pi$, $u(X) = u(r, \phi) = r^{\pi/\gamma}\sin(\phi\pi/\gamma)$ is a classical solution, and the approximation of $u(X)$ by finite element method is easy

(2) When $\pi < \gamma \leq 2\pi$, $u(X) = u(r, \phi) = r^{\pi/\gamma}\sin(\phi\pi/\gamma)$ is not a classical solution, and the approximation of $u(X)$ by finite element method is challenging. This BVP is one of the benchmark/test problem for developing new and better finite element method

The 1D time dependent heat equation: Again, consider a metal rod of length 1 occupying the interval $[0, 1]$. Let
$f(x, t)$ be the heat source intensity, $q(x, t)$ be the heat flux
$e(x, t)$ be the energy density
$T(x, t)$ be the temperature distribution at the location x and time t

Conservation of energy: the rate of change of internal energy equals the sum of net heat flux and produced heat, i.e.,

$$\int_x^{x+h} \frac{\partial e(s,t)}{\partial t} ds = A(x)q(x,t) - A(x+h)q(x+h,t) + \int_x^{x+h} f(s,t)ds$$

which leads to a differential equation $\frac{\partial e}{\partial t} + \frac{\partial (Aq)}{\partial x} = f$
Fourie's law: $q = -k\frac{\partial T}{\partial x}$
A constitutive law: $e = cT$

These three laws together yield the 1D heat equation:

$$c\frac{\partial T}{\partial t} - (AkT_x)_x = f$$

Notation: for $G = G(x, t)$, we have $\frac{\partial G}{\partial x} = G_x$, $\frac{\partial^2 G}{\partial x^2} = G_{xx}$ but $\frac{\partial^{200} G}{\partial x^{200}} = G_{???}$

An initial boundary value problem (IBVP) for the heat equation: Find $T(x,t)$ such that

$$\text{the PDE}: \quad c\frac{\partial T(x,t)}{\partial t} - (AkT_x(x,t))_x = f,\ (x,t) \in (0,1) \times (0, t_{end}]$$
$$\text{the boundary conditions}: \quad T(0,t) = g_0(t), T(1,t) = g_1(t),\ t \in (0, t_{end}]$$
$$\text{the initial condition}: \quad T(x,0) = T_0(x),\ x \in (0,1)$$

Of course we also consider the (IBVP) for the heat equation in higher dimesion:

$$\text{the PDE}: \quad c\frac{\partial T(X,t)}{\partial t} - \nabla \cdot (Ak\nabla T(X,t)) = f(X,t),\ (X,t) \in \Omega \times (0, t_{end}]$$
$$\text{the boundary conditions}: \quad T(X,t) = g(X,t),\ (X,t) \in \partial\Omega \times (0, t_{end}]$$
$$\text{the initial condition}: \quad T(X,0) = T_0(X),\ X \in \Omega$$

where $X = (x,y) \in \mathbb{R}^2$ or $X = (x,y,z) \in \mathbb{R}^3$.

An IBVP for the acoustic wave equation: Consider a domain Ω occupied by a liquid or gas. Consider the following quantities:
ρ is the density
$p(X,t)$ is the pressure at the location X at the time t
u is the velocity.

Newton's 2nd law: $\rho\frac{\partial u}{\partial t} = -\nabla p$
Conservation of energy: $\frac{\partial p}{\partial t} = -k\nabla u$

The acoustic wave equation: $\frac{\partial^2 p}{\partial t^2} = c^2\Delta p$, where $c^2 = k/\rho$.

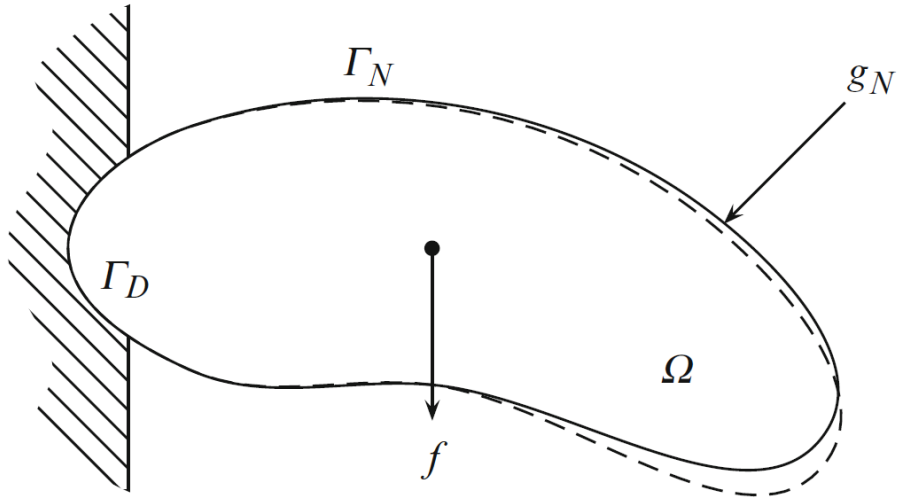
An IBVP for the acoustic wave equation:

$$\text{the PDE :} \quad \frac{\partial^2 p(X,t)}{\partial t^2} - c^2\Delta p(X,t) = f(X,t),\ (X,t) \in \Omega \times (0, t_{end})$$
$$\text{the boundary conditions :} \quad p|_{\partial\Omega} = g(X,t),\ (X,t) \in \partial\Omega \times (0, t_{end})$$
$$\text{the initial condition :} \quad p(X,0) = u_0(X),\ \dot{p}(X,0) = v_0(X),\ X \in \Omega$$

Notation:

$$\frac{\partial p}{\partial t} = \dot{p} = p_t, \quad \frac{\partial^2 p}{\partial t^2} = \ddot{p} = p_{tt}$$

A linear elatostatic problem: Consider a volume $\Omega \subset \mathbb{R}^3$ occupied by an elastic material with boundary $\partial\Omega$.



Let ω be an arbitrary subdomain of Ω. On ω, we consider the following forces:
$\mathbf{f}$ is the body force acting on the whole volume of ω.
$\sigma \cdot n$ is the surface force due to the stress σ which is a second order symmetric tensor.

Total force on ω: $\mathbf{F} = \int_\omega \mathbf{f} dX + \int_{\partial\omega} \sigma \cdot \mathbf{n} ds$
Applying the divergence theorem, we have

$$\mathbf{F} = \int_\omega (\mathbf{f} + \nabla \cdot \sigma) dX$$

In equilibrium, we should have $\mathbf{F} = \mathbf{0} \Rightarrow \ \mathbf{f} + \nabla \cdot \sigma = \mathbf{0}$

Note that $\nabla \cdot \sigma$ is the divergence of the stress tensor σ

$$\sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix}$$

Cauchy's equilibrium equation: $\mathbf{f} + \nabla \cdot \sigma = \mathbf{0}$.

$$f_1 + \frac{\partial \sigma_{11}}{\partial x_1} + \frac{\partial \sigma_{21}}{\partial x_2} + \frac{\partial \sigma_{31}}{\partial x_3} = 0 \tag{11.4a}$$

$$f_2 + \frac{\partial \sigma_{12}}{\partial x_1} + \frac{\partial \sigma_{22}}{\partial x_2} + \frac{\partial \sigma_{32}}{\partial x_3} = 0 \tag{11.4b}$$

$$f_3 + \frac{\partial \sigma_{13}}{\partial x_1} + \frac{\partial \sigma_{23}}{\partial x_2} + \frac{\partial \sigma_{33}}{\partial x_3} = 0 \tag{11.4c}$$

which look different from (11.4a)-(11.4c) in Book 2, but they are actually the same because of the symmetry of stress tensor σ.

Note that there are six unknowns in Cauchy's equation in these three equations:

$$\sigma_{11}, \sigma_{12}, \sigma_{13}, \sigma_{22}, \sigma_{23}, \sigma_{33}$$

Hence, we need to introduce constitutive equations to close the system of PDEs.

The constitutive equation is based on the relation between the stress and the strain. The strain tensor is defined according to the displacement $\mathbf{u}$:

$$\epsilon = \frac{1}{2}\left(\nabla \mathbf{u} + \nabla \mathbf{u}^T\right) \tag{11.5}$$

$$\epsilon_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right), \quad i,j = 1,2,3 \tag{11.6}$$

where $X = (x, y, z) = (x_1, x_2, x_3)$.

Hook's law for linear isotropic elastic materials:

$$\sigma = 2\mu\epsilon(\mathbf{u}) + \lambda(\nabla \cdot \mathbf{u})I \tag{11.7}$$

where I is the 3×3 identity matrix. The elastic moduli μ and λ are the so called Lamé parameters, defined

$$\mu = \frac{E}{2(1+\nu)}, \quad \lambda = \frac{E\nu}{(1+\nu)(1-2\nu)} \tag{11.8}$$

where E is Young's elastic modulus, and ν is Poisson's ratio.

A BVP for linear elastostatics: Find σ and $\mathbf{u}$ such that

$$-\nabla \cdot \sigma = f \tag{11.9a}$$

$$\sigma = 2\mu\epsilon(\mathbf{u}) + \lambda(\nabla \cdot \mathbf{u})I \tag{11.9b}$$

$$\mathbf{u}|_{\Gamma_D} = \mathbf{g}_D \tag{11.9c}$$

$$\sigma \cdot \mathbf{n}|_{\Gamma_N} = \mathbf{g}_N \tag{11.9d}$$

where $\Gamma_D \cup \Gamma_N = \partial\Omega$.

Of course we can substitute the formula of σ given in the second PDE into the first one to have a PDE about the displacement $\mathbf{u}$, and this leads to a BVP for $\mathbf{u}$ with the same boundary conditions. This is a HW problem.

Similar BVP can be considered for a two dimensional linear elastostatic object $\Omega \subset \mathbb{R}^2$.

Fluid Mechanics: Consider a domain Ω occupied by a fluid, let

ρ be the density, μ be the viscosity, $\mathbf{u}$ be the flow velocity
σ is the stress tensor of the fluid, p is the pressure

Conservation of mass: $$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0 \tag{12.2}$$

When ρ is a constant, this reduces to

$$\nabla \cdot \mathbf{u} = 0 \tag{12.3}$$

Momentum balance: $$\rho \frac{\partial u}{\partial t} + \rho(\mathbf{u} \cdot \nabla)\mathbf{u} = \nabla \cdot \sigma + \mathbf{f} \tag{12.8}$$

Newtonian fluid: $$\sigma = -pI + \mu(\nabla \mathbf{u} + \nabla \mathbf{u}^T) \tag{12.10}$$

Navier-Stokes equations for incompressible Newtonian fluids:

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot grad)\mathbf{u} = \nu \Delta \mathbf{u} - \frac{1}{\rho}\nabla p + \mathbf{f} \tag{12.11a}$$

$$\nabla \cdot \mathbf{u} = 0 \tag{12.11b}$$

Here $\nu = \mu/\rho$ is the kinematic viscosity.

A BVP for Stokes system: For a laminar stationary fluid, the Navier-Stokes system reduces to the Stokes system. We can then consider the following BVP

$$-\nu\Delta\mathbf{u} + \frac{1}{\rho}\nabla p = \mathbf{f} \text{ in } \Omega \tag{12.12a}$$

$$\nabla \cdot \mathbf{u} = 0 \text{ in } \Omega \tag{12.12b}$$

$$\mathbf{u} = \mathbf{g}_D \text{ on } \partial\Omega \tag{12.12c}$$

We have surveyed a variety (eight??) problems for partial differential equations (PDEs). Even though their underlying theories are diversified, finite element method is a computational technique that can provide approximate solutions to PDEs in various types.

One Problem, Different Formulations: Given a nonsingular $n \times n$ matrix K and vector $\mathbf{f} \in \mathbb{R}^n$. Consider

Problem E (Equation Problem): Find $\mathbf{u} \in \mathbb{R}^n$ such that $K\mathbf{u} = \mathbf{f}$.

Problem G (Galerkin Problem): Find $\mathbf{u} \in \mathbb{R}^n$ such that

$$\mathbf{v} \cdot K\mathbf{u} = \mathbf{v} \cdot \mathbf{f}, \quad \forall \mathbf{v} \in \mathbb{R}^n$$

Problem R (Riesz Problem): Assume that there is a 2-times continuously differentiable function $F(\mathbf{v})$ such that

$$\nabla F(\mathbf{v}) = K\mathbf{v} - \mathbf{f}$$

Find a minimizer $\mathbf{u} \in \mathbb{R}^n$ for $F(\mathbf{v})$.

Note that the existence of F that is 2-times continuously differentiable and $\nabla F(\mathbf{v}) = K\mathbf{v} - \mathbf{f}$ implies that K symmetric and

$$F(\mathbf{v}) = \frac{1}{2}\mathbf{v} \cdot K\mathbf{v} - \mathbf{v} \cdot \mathbf{f}$$

(HW???)
Problem L Problem (Least Squares): Find $\mathbf{u}$ that minimizes

$$L(\mathbf{v}) = \|K\mathbf{v} - \mathbf{f}\|^2$$

With appropriate assumptions, all of these problems have the same solution because we can establish the following equivalence:

$$\begin{aligned} \textbf{Problem E} &\iff \textbf{Problem G} \\ \textbf{Problem E} &\iff \textbf{Problem R} \\ \textbf{Problem E} &\iff \textbf{Problem L} \end{aligned}$$

For simplicity, we consider an SPD matrix $K = (k_{ij})_{i,j=1}^{2}$ and a vector $\mathbf{f} = (f_i)_{i=1}^{2}$.

Assume $\mathbf{u} = (u_1, u_2)^T$ is the solution to $K\mathbf{u} = \mathbf{f}$. Then

$$\begin{aligned} k_{11}u_1 + k_{12}u_2 &= f_1 \\ k_{21}u_1 + k_{22}u_2 &= f_2 \end{aligned}$$

Then the solution $\mathbf{u}$ to **Problem E** is such that

$$\begin{aligned} & v_1(k_{11}u_1 + k_{12}u_2 - f_1) + v_2(k_{21}u_1 + k_{22}u_2 - f_2) = 0 \quad \text{for any } v_1, v_2 \\ & \mathbf{v} \cdot K\mathbf{u} = \mathbf{v} \cdot \mathbf{f}, \quad \mathbf{v} \in \mathbb{R}^2 \end{aligned}$$

i.e., $\mathbf{u}$ is a solution to **Problem G**. This proves **Problem E** $\Longrightarrow$ **Problem G**.

Conversely, assume that $\mathbf{u}$ is a solution to **Problem G**. Then

$$\mathbf{v} \cdot K\mathbf{u} = \mathbf{v} \cdot \mathbf{f}, \quad \mathbf{v} \in \mathbb{R}^2$$
$$v_1(k_{11}u_1 + k_{12}u_2 - f_1) + v_2(k_{21}u_1 + k_{22}u_2 - f_2) = 0 \quad \text{for any } v_1, v_2$$

Letting $v_1 = 1, v_2 = 0$ and letting $v_1 = 0, v_2 = 1$ lead to

$$\begin{aligned} k_{11}u_1 + k_{12}u_2 &= f_1 \\ k_{21}u_1 + k_{22}u_2 &= f_2 \end{aligned}$$

or $K\mathbf{u} = \mathbf{u}$ which implies that $\mathbf{u}$ is a solution to **Problem E**.

Combining the above two arguments together, we have the following equivalence:

$$\textbf{Problem E} \iff \textbf{Problem G}$$

For **Problem E** $\Longleftrightarrow$ **Problem R**: Recall

$$F(\mathbf{v}) = \frac{1}{2}\mathbf{v} \cdot K\mathbf{v} - \mathbf{v} \cdot \mathbf{f}, \quad \nabla F(\mathbf{v}) = K\mathbf{v} - \mathbf{f}$$

First, assume that $\mathbf{u} = (u_1, u_2)^T$ solves **Problem E**, i.e., $\mathbf{u}$ is the solution to $K\mathbf{u} = \mathbf{f}$. Then

$$\nabla F(\mathbf{u}) = K\mathbf{u} - \mathbf{f} = \mathbf{0}$$

meaning that $\mathbf{u}$ is a critical point for function $F(\mathbf{v})$. Moreover, the Hessian of $F(\mathbf{v})$ is K which is SPD; hence, function $F(\mathbf{v})$ reaches its minimum at $\mathbf{u}$, i.e., $\mathbf{u}$ is the minimizer of $F(\mathbf{v})$ or $\mathbf{u}$ solves **Problem R**. This proves **Problem E** $\Longrightarrow$ **Problem R**.

Conversely, assume that $\mathbf{u}$ solves **Problem R**, i.e., $\mathbf{u}$ is the minimizer of $F(\mathbf{v})$. Then,

$$K\mathbf{u} - \mathbf{f} = \nabla F(\mathbf{u}) = \mathbf{0}$$

so $K\mathbf{u} = \mathbf{f}$ meaning $\mathbf{u}$ solves **Problem E**. This proves **Problem R** $\Longrightarrow$ **Problem E**.

The combination of these two arguments leads to **Problem E** $\Longleftrightarrow$ **Problem R**.

Similarly, we can show **Problem E** $\Longleftrightarrow$ **Problem L** (HW???).

Recall: Given a SPD matrix K and a vector $\mathbf{f}$

Problem E (Equation Problem): Find $\mathbf{u} \in \mathbb{R}^n$ such that $K\mathbf{u} = \mathbf{f}$.

Problem G (Galerkin Problem): Find $\mathbf{u} \in \mathbb{R}^n$ such that

$$\mathbf{v} \cdot K\mathbf{u} = \mathbf{v} \cdot \mathbf{f}, \quad \forall \mathbf{v} \in \mathbb{R}^n$$

Problem R (Riesz Problem): For

$$F(\mathbf{v}) = \frac{1}{2}\mathbf{v} \cdot K\mathbf{v} - \mathbf{v} \cdot \mathbf{f}$$

Find a minimizer $\mathbf{u} \in \mathbb{R}^n$ for $F(\mathbf{v})$.

Problem L Problem (Least Squares): Find $\mathbf{u}$ that minimizes

$$L(\mathbf{v}) = \|K\mathbf{v} - \mathbf{f}\|^2$$

Problems G, R, L all seem to be more complicated than **Problem E**, but they provide alternative approaches for solving **Problem E** that maybe be more efficiently on computers.

Finite element methods for 1D problems:

A model boundary values problems

Consider the following general 1D elliptic BVP: find $u(x)$ such that

$$-(au')' + cu = f, \ x \in (0, L) \tag{2.40a}$$
$$au'(0) = k_0(u(0) - g_0) \tag{2.40b}$$
$$-au'(L) = k_L(u(L) - g_L) \tag{2.40c}$$

where functions $a > 0$, $c \geq 0$, and f are given functions, and $k_0 \geq 0, k_L \geq 0, g_0$ and g_L are given parameters.

Boundary condition for $k_0 = \infty$ or $k_0 = 0$ leads to the Dirichlet or Neumann boundary condition

Boundary condition for $k_L = \infty$ or $k_L = 0$ leads to the Dirichlet or Neumann boundary condition

We start from the following model 1D elliptic BVP: find $u(x)$ such that

$$-(au')' + cu = f, \ x \in (0,1) \tag{1}$$
$$a(0)u'(0) = g_0 \tag{2}$$
$$u(1) = g_1 \tag{3}$$

Derivation of the weak equation: multiply equation (1) by a test function v, integrate both sides, and apply integration by parts:

$$-v(au')' + cvu = vf$$
$$-\int_0^1 v(au')'dx + \int_0^1 cvudx = \int_0^1 vfdx$$
$$\int_0^1 av'u'dx + v(0)a(0)u'(0) - v(1)a(1)u'(1) + \int_0^1 cvudx = \int_0^1 vfdx$$

Then, applying the boundary condition, we have the weak equation for the given differential equation:

$$\int_0^1 av'u'dx + v(0)g_0 - v(1)a(1)u'(1) + \int_0^1 cvudx = \int_0^1 vfdx \tag{4}$$

Identify the test and trial function sets: Recall that the previous weak equation is

$$\int_0^1 av'u'dx - v(1)a(1)u'(1) + \int_0^1 cvudx = \int_0^1 vfdx - v(0)g_0 \tag{4}$$

Since our boundary condition does not provide any information about $a(1)u'(1)$, we would like to avoid it in the weak equation. This can be done by choosing a test function v such that $v(1) = 0$. Hence the weak equation becomes

$$\int_0^1 av'u'dx + \int_0^1 cvudx = \int_0^1 vfdx - v(0)g_0 \tag{5}$$

This weak equation suggests the following

(1). the test function v can be chose from the following function space

$$H^1(0,1) = \{w \mid w \in L^2(0,1), w' \in L^2(0,1)\} \tag{6}$$

and avoiding/eliminating $a(1)u'(1)$ further suggests choosing the test function v from

$$\mathcal{T} = \{w \mid w \in H^1(0,1), w(1) = 0\} \tag{7}$$

(2). the trial function u can be chosen from the following SET of functions:

$$\mathcal{S} = \{w \mid w \in H^1(0,1), w(1) = g_1\} \tag{8}$$

For the model 1D elliptic BVP

$$-(au')' + cu = f, \ x \in (0,1) \tag{1}$$
$$a(0)u'(0) = g_0 \tag{2}$$
$$u(1) = g_1 \tag{3}$$

A Weak Form for this BVP: find $u \in \mathcal{S}$ such that

$$\int_0^1 av'u'dx + \int_0^1 cvudx = \int_0^1 vfdx - v(0)g_0, \ \forall v \in \mathcal{T} \tag{5}$$

where
$$\mathcal{T} = \{w \mid w \in H^1(0,1), w(1) = 0\} \tag{7}$$
$$\mathcal{S} = \{w \mid w \in H^1(0,1), w(1) = g_1\} \tag{8}$$

Remarks:

- A weak form of a BVP has three components: the weak equation, the test function space, and the trial function set.
- A boundary condition in a BVP is natural if it is applicable "naturally" in the weak equation. A boundary condition is essential when it has to be enforced through suitable requirements on the test function space and trial function set. Boundary condition (2) is natural, but (3) is essential.
- Does the BVP have a solution? We will not discuss this b/c .....
  Does the weak problem have a solution? A unique solution? The general answer is yes according to the Lax-Milgram theorem to be discussed later.

The Petrov-Galerkin method for the model 1D elliptic BVP: Choose two parameters m and n and some functions $\phi_i, 1 \le i \le m, \psi_j, 1 \le j \le n$ such that

$$\begin{aligned} &\text{A set of test functions :} && \mathcal{T}_n = span\{\psi_1(x), \psi_2(x), \cdots, \psi_n(x)\} \subseteq \mathcal{T} \\ &\text{A set of trial functions :} && \mathcal{S}_m = span\{\phi_1(x), \phi_2(x), \cdots, \phi_m(x)\} \end{aligned}$$

By $\mathcal{T}_n \subseteq \mathcal{T}$, we have $\psi_i(1) = 0, 1 \le i \le n$. We look for $u_{pg} \in \mathcal{S}_m$ such that

$$\begin{aligned} u_{pg}(1) &= g_1, \\ \int_0^1 a\psi_i' u_{pg}' dx + \int_0^1 c\psi_i u_{pg} dx &= \int_0^1 \psi_i f dx - \psi_i(0) g_0, \quad i = 1, 2, \cdots, n \end{aligned}$$

Because of $u_{pg}(x) = \sum_{j=1}^m u_j \phi_j(x)$, these equations can be written as

$$\begin{aligned} &\sum_{j=1}^m u_j \phi_j(1) = g_1 \\ &\sum_{j=1}^m \left(\int_0^1 a\psi_i' \phi_j' dx \right) u_j + \sum_{j=1}^m \left(\int_0^1 c\psi_i \phi_j dx \right) u_j = \int_0^1 \psi_i f dx - \psi_i(0) g_0 \\ &\qquad i = 1, 2, \cdots, n \end{aligned}$$

Recall: BVP $\Rightarrow$ weak problem (1 equation, 2 groups of functions):

P-G method: compute $u_{pg}(x) = \sum_{j=1}^{m} u_j \phi_j(x)$ with functions chosen in $\mathcal{T}_n, \mathcal{S}_m$ by

$$\sum_{j=1}^{m} u_j \phi_j(1) = g_1$$

$$\sum_{j=1}^{m} \left(\int_0^1 a \psi_i' \phi_j' dx \right) u_j + \sum_{j=1}^{m} \left(\int_0^1 c \psi_i \phi_j dx \right) u_j = \int_0^1 \psi_i f dx - \psi_i(0) g_0,$$

$$i = 1, 2, \cdots, n.$$

The matrix form of the equations for computing $\vec{u} = (u_j)_{j=1}^m$: $\tilde{M}\vec{u} = \tilde{\vec{f}} + \tilde{\vec{B}}$ with

$$\tilde{M}(1, 1:m) = [\phi_1(1), \phi_2(1), \cdots, \phi_m(1)], \quad \tilde{\vec{f}}(1) = g_1, \ \tilde{\vec{B}}(1) = 0$$

$$\tilde{M}(2: n+1, 1:m) = K + M$$

$$K = \left(\int_0^1 a \psi_i'(x) \phi_j'(x) dx \right)_{i=1,j=1}^{n,m}, \quad M = \left(\int_0^1 c \psi_i(x) \phi_j(x) dx \right)_{i=1,j=1}^{n,m}$$

$$\tilde{\vec{f}}(2: n+1) = \vec{f} = \left(\int_0^1 \psi_j(x) f(x) dx \right)_{i=1}^{n}$$

$$\tilde{\vec{B}}(2: n+1) = -g_0 (\psi_i(0))_{i=1}^n$$

Example: Consider the BVP with $a(x) = 1, c(x) = 0$

$$-u'' = \sin(x-1), \quad x \in (0,1)$$
$$u'(0) = (0.2 + cos(1)), \quad u(1) = 0.3$$

Recall the weak equation:

$$\int_0^1 av'u'dx + \int_0^1 cvudx = \int_0^1 vfdx - v(0)g_0, \ \forall v \in \mathcal{T} \tag{5}$$

To apply the PG method with the following test and trial function space:

$$\begin{aligned} \mathcal{T}_2 &= span\{\psi_1(x), \psi_2(x)\} = span\{\sin(x-1), (x-1)\} \\ \mathcal{S}_3 &= span\{\phi_1(x), \phi_2(x), \phi_3(x)\} = span\{\cos(x), e^x, x\} \end{aligned}$$

Then we look for a PG solution $u_{pg}(x) = u_1\phi_1(x) + u_2\phi_2(x) + u_3\phi_3(x)$ such that

$$u_1\phi_1(x) + u_2\phi_2(x) + u_3\phi_3(x) = u_{pg}(1) = g_1 \quad \text{(the essential boundary condition)}$$
$$\int_0^1 \psi_1'(x)u_{pg}'(x)dx = \int_0^1 \psi_1(x)f(x)dx - \psi_1(0)g_0$$
$$\int_0^1 \psi_2'(x)u_{pg}'(x)dx = \int_0^1 \psi_2(x)f(x)dx - \psi_2(0)g_0$$

with $f(x) = \sin(x-1), \ g_0 = (0.2 + \cos(1)), \ g_1 = 0.3.$

Recall the P-G equations:

$$u_1\phi_1(x) + u_2\phi_2(x) + u_3\phi_3(x) = u_{pg}(1) = g_1 \quad \text{(the essential boundary condition)}$$
$$\int_0^1 \psi_1'(x)u_{pg}'(x)dx = \int_0^1 \psi_1(x)f(x)dx - \psi_1(0)g_0$$
$$\int_0^1 \psi_2'(x)u_{pg}'(x)dx = \int_0^1 \psi_2(x)f(x)dx - \psi_2(0)g_0$$

and $u_{pg}(x) = u_1\phi_1(x) + u_2\phi_2(x) + u_3\phi_3(x)$, we can reduce them to

$$u_1\phi_1(x) + u_2\phi_2(x) + u_3\phi_3(x) = u_{pg}(1) = g_1 \quad \text{(the essential boundary condition)}$$
$$\int_0^1 \psi_1'(x)(u_1\phi_1'(x) + u_2\phi_2'(x) + u_3\phi_3'(x))dx = \int_0^1 \psi_1(x)f(x)dx - \psi_1(0)g_0$$
$$\int_0^1 \psi_2'(x)(u_1\phi_1'(x) + u_2\phi_2'(x) + u_3\phi_3'(x))dx = \int_0^1 \psi_2(x)f(x)dx - \psi_2(0)g_0$$

In matrix form, we have

$$\begin{bmatrix} \phi_1(1) & \phi_2(1) & \phi_3(1) \\ \int_0^1 \psi_1'(x)\phi_1'(x)dx & \int_0^1 \psi_1'(x)\phi_2'(x)dx & \int_0^1 \psi_1'(x)\phi_3'(x)dx \\ \int_0^1 \psi_2'(x)\phi_1'(x)dx & \int_0^1 \psi_2'(x)\phi_2'(x)dx & \int_0^1 \psi_2'(x)\phi_3'(x)dx \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$
$$= \begin{bmatrix} 0.3 \\ \int_0^1 \psi_1(x)f(x)dx - \psi_1(0)g_0 \\ \int_0^1 \psi_2(x)f(x)dx - \psi_2(0)g_0 \end{bmatrix} = \begin{bmatrix} 0 \\ \int_0^1 \psi_1(x)f(x)dx \\ \int_0^1 \psi_2(x)f(x)dx \end{bmatrix} + \begin{bmatrix} 0.3 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -\psi_1(0)g_0 \\ -\psi_2(0)g_0 \end{bmatrix}$$

We can develop a Matlab script for solving for $\mathbf{u} = [u_1; u_2; u_3]$ from this linear system as follows.

```
clear;
ph1 = @(x) cos(x); ph2 = @(x) exp(x); ph3 = @(x) x;
ph1d = @(x) -sin(x); ph2d = @(x) exp(x); ph3d = @(x) ones(size(x));
ps1 = @(x) sin(x-1); ps2 = @(x) x-1;
ps1d = @(x) cos(x-1); ps2d = @(x) ones(size(x));

M = zeros(3,3);
M(1, :) = [ph1(1), ph2(1), ph3(1)];
tmp = @(x) ps1d(x).*ph1d(x); M(2,1)  = integral(tmp, 0, 1);
tmp = @(x) ps1d(x).*ph2d(x); M(2,2)  = integral(tmp, 0, 1);
tmp = @(x) ps1d(x).*ph3d(x); M(2,3)  = integral(tmp, 0, 1);
tmp = @(x) ps2d(x).*ph1d(x); M(3,1)  = integral(tmp, 0, 1);
tmp = @(x) ps2d(x).*ph2d(x); M(3,2)  = integral(tmp, 0, 1);
tmp = @(x) ps2d(x).*ph3d(x); M(3,3)  = integral(tmp, 0, 1);

f = @(x) sin(x-1); rhsf = zeros(3,1);
tmp = @(x) ps1(x).*f(x); rhsf(2) = integral(tmp,0,1);
tmp = @(x) ps2(x).*f(x); rhsf(3) = integral(tmp,0,1);
nbc = 0.2 + cos(1);
rhs = rhsf + [0.3; -ps1(0)*nbc; -ps2(0)*nbc];
vu = M\rhs;
```

Remark: Matlab's function `integral` is inefficient, and we should not use it in our finite element codes.

This Matlab script produces a vector:

```
vu =  -0.681169405799283
      -0.060301579008613
       0.831954087086684
```

which can be used to form a function $u_{pg}(x)$ as the approximation to the exact solution $u(x) = \sin(x-1) + 0.2(x-1) + 0.3$ for the BVP:

```
u_pg = @(x) vu(1)*ph1(x) + vu(2)*ph2(x) + vu(3)*ph3(x);
u_true = @(x) sin(x-1) + 0.2*(x-1) + 0.3;
x = 0:0.01:1;
plot(x, u_pg(x), 'r.', x, u_true(x), 'b')
xlabel('x')
title('A simple Petrov-Galerkin solution')
legend('PG solution u_{pg}(x)', 'true solution u(x)', ...
       'Location', 'NorthWest')
```

Computationally, we compute with this Matlab script/program for generating numbers, but conceptually/mathematically, we obtain a function $u_{pg}(x)$ that can approximate the exact solution $u(x)$ for a BVP.

This is quite different from the Finite Difference method for solving differential equations in which there is no underlying approximation function, at least not explicitly.

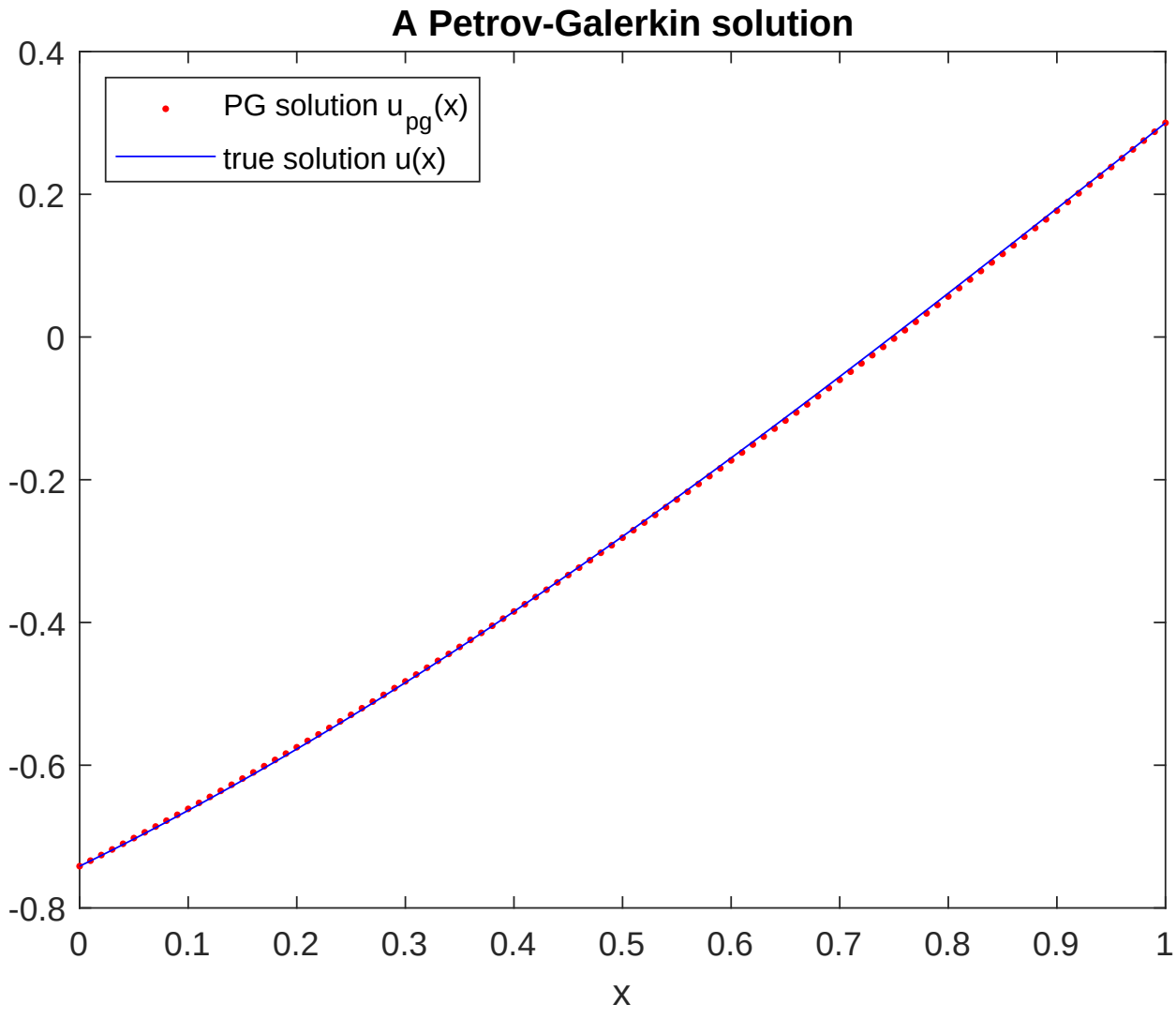


Figure: the exact and the PG solutions for the model 1D elliptic BVP.

The Bubnov-Galerkin method: Recall the model 1D elliptic BVP is to find u such that

$$-(au')' + cu = f, \ x \in (0,1) \tag{1}$$
$$a(0)u'(0) = g_0 \tag{2}$$
$$u(1) = g_1 \tag{3}$$

A Weak Form for this BVP: find $u \in \mathcal{S}$ such that

$$\int_0^1 av'u'dx + \int_0^1 cvudx = \int_0^1 vfdx - v(0)g_0, \ \forall v \in \mathcal{T} \tag{5}$$

where

$$\mathcal{T} = \{w \mid w \in H^1(0,1), w(1) = 0\} \tag{7}$$
$$\mathcal{S} = \{w \mid w \in H^1(0,1), w(1) = g_1\} \tag{8}$$

Instead of what used in the P-G method:

$$\mathcal{S}_m = span\{\phi_1(x), \phi_2(x), \cdots, \phi_m(x)\}$$
$$\mathcal{T}_n = span\{\psi_1(x), \psi_2(x), \cdots, \psi_n(x)\} \subseteq \mathcal{T}$$

now, we choose

$$\mathcal{T}_n = span\{\psi_1(x), \psi_2(x), \cdots, \psi_n(x)\} \subseteq \mathcal{T}$$
$$\mathcal{S}_n = \left\{ \sum_{j=1}^{n} u_j\psi_j(x) + G(x), \quad u_j \in \mathbb{R}^1 \right\}$$

with $G(1) = g_1$. Then the Petrov-Galerkin method becomes the so called Bubnov-Galerkin or simply Galerkin method. In this case, since $\mathcal{T}_n \subseteq \mathcal{T}$, we must have $\psi_j(1) = 0, \ j = 1, 2, \cdots, n$

The Galerkin solution is in the following form $u_g(x) = \sum_{j=1}^{n} u_j \psi_j(x) + G(x)$.
The Galerkin equation for computing $\vec{u} = (u_j)_{j=1}^{n}$ can be derived as follows:

$$\int_0^1 av'u'dx + \int_0^1 cvudx = \int_0^1 vfdx - v(0)g_0, \ \forall v \in \mathcal{T} \tag{5}$$

$$\begin{aligned} & \sum_{j=1}^{n} \left(\int_0^1 a\psi_i'\psi_j'dx \right) u_j + \sum_{j=1}^{n} \left(\int_0^1 c\psi_i\psi_j dx \right) u_j \\ = & \int_0^1 \psi_i f dx - \psi_i(0)g_0 - \int_0^1 a\psi_i' G'(x)dx - \int_0^1 c\psi_i G(x)dx, \quad i = 1, 2, \cdots, n \end{aligned}$$

Again, we can write these equations about $\vec{u} = (u_j)_{j=1}^{n}$ in matrix form as $(K + M)\vec{u} = \vec{f} + \vec{B}$ with

$$K = \left(\int_0^1 a\psi_i'(x)\psi_j'(x)dx \right)_{i=1,j=1}^{n,n}, \quad M = \left(\int_0^1 c\psi_i(x)\psi_j(x)dx \right)_{i=1,j=1}^{n,n}$$

$$\vec{f} = \left(\int_0^1 \psi_j(x)f(x)dx \right)_{i=1}^{n}$$

$$\vec{B} = -g_0(\phi_i(0))_{i=1}^{n} - \left(\int_0^1 a\psi_i' G'(x)dx \right)_{i=1}^{n} - \left(\int_0^1 c\psi_i G(x)dx \right)_{i=1}^{n}$$

Remark: each integral on the left of the weak equation leads a matrix.

Example: Consider the BVP with $a(x) = 1, c(x) = 0$

$$-u'' = sin(x-1), \quad x \in (0,1),$$
$$u'(0) = (0.2 + \cos(1)), \quad u(1) = 0.3$$

To apply the Galerkin method, we choose

$$\begin{aligned} \mathcal{T}_2 &= span\{\psi_1(x), \psi_2(x)\} = span\{\sin(x-1), (x-1)\}, \\ \mathcal{S}_2 &= span\{\psi_1(x), \psi_2(x)\} + G(x) = span\{\sin(x-1), (x-1)\} + G(x), \quad G(1) = 0.3 \end{aligned}$$

For example, we can choose $G(x) = 0.3x$. Then we look for a Galerkin solution

$$u_g(x) = u_1\psi_1(x) + u_2\psi_2(x) + G(x) = u_1 \sin(x-1) + u_2(x-1) + 0.3x$$

By the weak equation:

$$\int_0^1 av'u'dx + \int_0^1 cvudx = \int_0^1 vfdx - v(0)g_0, \ \forall v \in \mathcal{T} \tag{5}$$

we have the Bubnov-Galerkin equations:

$$\int_0^1 \psi_1'(x)u_g'(x)dx = \int_0^1 \psi_1(x)f(x)dx - \psi_1(0)(0.2 + \cos(1))$$
$$\int_0^1 \psi_2'(x)u_g'(x)dx = \int_0^1 \psi_2(x)f(x)dx - \psi_2(0)(0.2 + \cos(1))$$

Recall the Bubnov-Galerkin equations:

$$\int_0^1 \psi_1'(x)u_g'(x)dx = \int_0^1 \psi_1(x)f(x)dx - \psi_1(0)(0.2 + cos(1))$$
$$\int_0^1 \psi_2'(x)u_g'(x)dx = \int_0^1 \psi_2(x)f(x)dx - \psi_2(0)(0.2 + cos(1))$$
$$\text{with} \quad u_g(x) = u_1\psi_1(x) + u_2\psi_2(x) + G(x) = u_1 \sin(x-1) + u_2(x-1) + 0.3x$$

we can reduce the Bubnov-Galerkin equations to

$$\int_0^1 \psi_1'(u_1\psi_1' + u_2\psi_2')dx = \int_0^1 \psi_1 f dx - \psi_1(0)(0.2 + cos(1)) - \int_0^1 \psi_1' G' dx$$
$$\int_0^1 \psi_2'(u_1\psi_1' + u_2\psi_2')dx = \int_0^1 \psi_2 f dx - \psi_2(0)(0.2 + cos(1)) - \int_0^1 \psi_2' G' dx$$

The matrix form the Galerkin equations is

$$\begin{bmatrix} \int_0^1 \psi_1'\psi_1' dx & \int_0^1 \psi_1'\psi_2' dx \\ \int_0^1 \psi_2'\psi_1' dx & \int_0^1 \psi_2'\psi_2' dx \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$
$$= \begin{bmatrix} \int_0^1 \psi_1 f dx \\ \int_0^1 \psi_2 f dx \end{bmatrix} - (0.2 + \cos(1)) \begin{bmatrix} \psi_1(0) \\ \psi_2(0) \end{bmatrix} - \begin{bmatrix} \int_0^1 \psi_1' G' dx \\ \int_0^1 \psi_2' G' dx \end{bmatrix}$$

Solving for $u_1 = 1, u_2 = -0.1$ and $u_g(x) = \sin(x-1) - 0.1(x-1) + 0.3x$ which is the exact solution!

A Matlab script for solving for **u** from the Bubnov-Galerkin equations:

```
clear;
ps1 = @(x) sin(x-1); ps2 = @(x) x-1;
ps1d = @(x) cos(x-1); ps2d = @(x) ones(size(x));
G = @(x) 0.3*x; Gd = @(x) 0.3*ones(size(x));

M = zeros(2,2);
tmp = @(x) ps1d(x).*ps1d(x); M(1,1)  = integral(tmp, 0, 1);
tmp = @(x) ps1d(x).*ps2d(x); M(1,2)  = integral(tmp, 0, 1);
tmp = @(x) ps2d(x).*ps1d(x); M(2,1)  = integral(tmp, 0, 1);
tmp = @(x) ps2d(x).*ps2d(x); M(2,2)  = integral(tmp, 0, 1);

f = @(x) sin(x-1); rhsf = zeros(2,1);
tmp = @(x) ps1(x).*f(x); rhsf(1) = integral(tmp,0,1);
tmp = @(x) ps2(x).*f(x); rhsf(2) = integral(tmp,0,1);
rhsN = -(0.2 + cos(1))*[ps1(0); ps2(0)]; % Natural BC
rhsE = zeros(2, 1); % Essential BC
tmp = @(x) ps1d(x).*Gd(x); rhsE(1) = -integral(tmp,0,1);
tmp = @(x) ps2d(x).*Gd(x); rhsE(2) = -integral(tmp,0,1);
rhs = rhsf + rhsE + rhsN;

vu = M\rhs;
```

The script above will produce

```
vu = 1.000000000000007
     -0.100000000000006
```

which essentially indicate `vu(1) = 1, vu(2) = -0.1`. A Matlab script for using $u_g(x)$:

```
u_g = @(x) vu(1)*ps1(x) + vu(2)*ps2(x) + G(x);
u_true = @(x) sin(x-1) + 0.2*(x-1) + 0.3;
x = 0:0.01:1;
plot(x, u_g(x), 'r.', x, u_true(x), 'b')
xlabel('x')
title('A simple Galerkin solution')
legend('BG solution u_{g}(x)', 'true solution u(x)', ...
        'Location', 'NorthWest')
```

It is easy to see that $u_g(x)$ is coincidentally the exact solution $u(x) = \sin(x-1) + 0.2(x-1) + 0.3$.

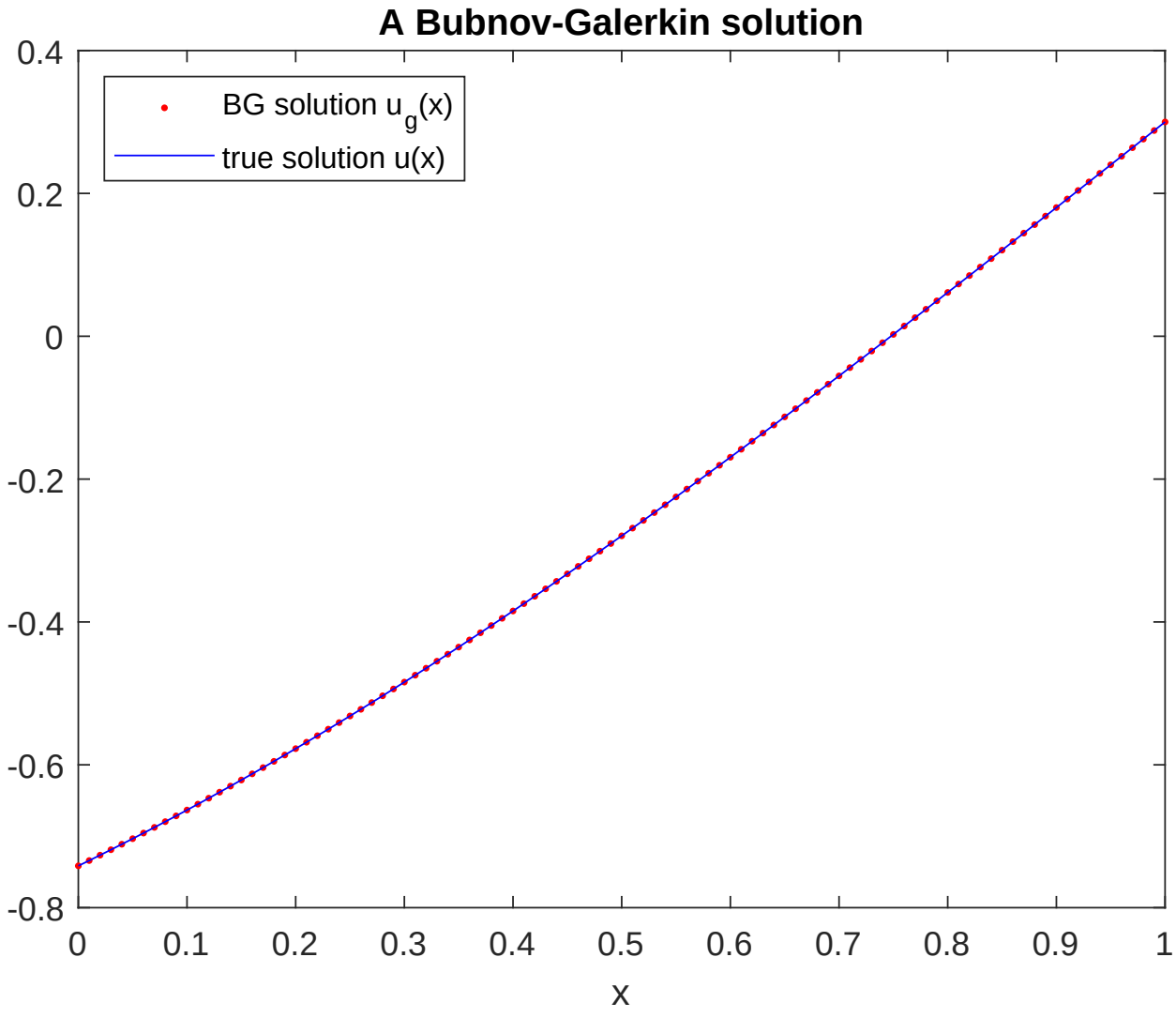


Figure: the exact and the Galerkin solutions for the model 1D elliptic BVP.

Remarks:

- In general $\mathcal{S}$ and $\mathcal{S}_m$ are not linear spaces because of the non-homogeneous boundary condition. However, we can try to homogenize the boundary condition and reduce the original problem to a BVP with homogeneous boundary condition. Then in the weak form of this new BVP, the solution set $\mathcal{S}$ can be chosen as a linear space.
- The dimensions of $\mathcal{S}_m$ and $\mathcal{T}_n$ do not have to be the same. In this situation, the number of equations and unknowns in the above linear system about the coefficients of u_{pg} are not the same, and special numerical methods have to be used to solve this linear system. We therefore usually use $\mathcal{S}_m$ and $\mathcal{T}_n$ such that $m = n$ in a Petrov-Galerkin method.
- We should choose $\phi_1(x), \phi_2(x), \cdots, \phi_n(x)$ and $\psi_1(x), \psi_2(x), \cdots, \psi_n(x)$ such that they are linearly independent.
- For a symmetric BVP, the matrix in the Galerkin method is symmetric, but the matrix in the PG method is not necessarily symmetric.

A generic framework to solve a BVP by the Galerkin method: Given a BVP such as

$$-(au')' + cu = f, \ x \in (0,1) \tag{1}$$
$$a(0)u'(0) = g_0 \tag{2}$$
$$u(1) = g_1 \tag{3}$$

(A). Derive A Weak Form for this BVP: find $u \in \mathcal{S}$ such that

$$\int_0^1 av'u'dx + \int_0^1 cvudx = \int_0^1 vfdx - v(0)g_0, \ \forall v \in \mathcal{T} \tag{5}$$

where $\mathcal{T} = \{w \mid w \in H^1(0,1), w(1) = 0\}$ (7)

$\mathcal{S} = \{w \mid w \in H^1(0,1), w(1) = g_1\}$ (8)

(B). Choose/construct $\mathcal{T}_n = span\{\psi_1(x), \psi_2(x), \cdots, \psi_n(x)\} \subseteq \mathcal{T}$, $G(x)$ and use the weak form to set up the Galerkin equation: $(K + M)\vec{u} = \vec{f} + \vec{B}$ which can produce a Galerkin solution

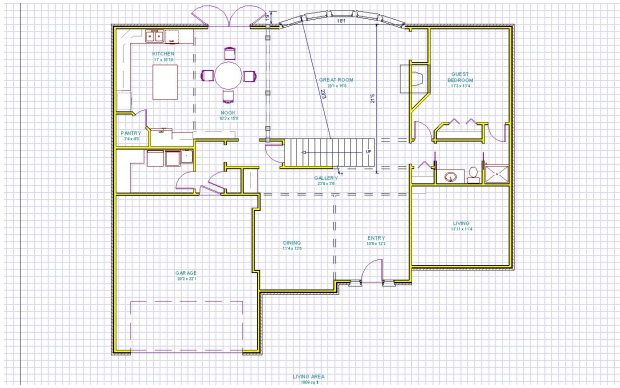
$$u_g(x) = \sum_{j=1}^{n} u_j\psi_j(x) + G(x) \approx u(x) = ?$$

A few questions: how to choose/create $\psi_i(x), \ i = 1, 2, \cdots, n$? How to find $G(x)$? How to form the matrices K and M? How to form the vectors $\vec{f}$ and $\vec{B}$?

The computations for $u_g(x)$ are essentially determined by our choice/creation of the functions $\psi_i(x), \ i = 1, 2, \cdots, n$.

A VERY sketchy framework for finite element computations

Construction:



Finite Element Methods:

Mesh, finite element functions/space, quadrature rules, etc

Use a weak form and a FE space to set up a Finite Element Scheme/Equation

Solve the FE equation to obtain a finite element solution