AME50541: Finite Element Methods
Polynomial Approximation

1 Polynomial spaces

Consider an open set $\Omega \subset \mathbb{R}^d$ and define $\mathbb{P}^p(\Omega)$ as the function space of polynomials of degree at most p over Ω, where the degree of a monomial of the form $x_1^{\alpha_1}\cdots x_d^{\alpha_d}$ for $x \in \Omega$, $\alpha \in \mathbb{N}^d$ is $\sum_{i=1}^d \alpha_i$. The space $\mathbb{P}^p(\Omega)$, spanned by the collection of all monomials up to and including degree p, is conveniently re-written as

$$\mathbb{P}^p(\Omega) := \operatorname{span}\left\{ f : \Omega \to \mathbb{R} \,\middle|\, f(x) = x_1^{\alpha_1}\cdots x_d^{\alpha_d},\, \alpha \in \mathbb{N}^d,\, 0 \leqslant \sum_{i=1}^d \alpha_i \leqslant p \right\}, \tag{1}$$

which emphasizes that $\mathbb{P}^p(\Omega)$ is a finite-dimensional vector space. Since the collection of monomials up to and including degree p is a basis for $\mathbb{P}^p(\Omega)$ (spans the whole space and all monomials are linearly independent), the dimension of $\mathbb{P}^p(\Omega)$ is equal to the number of monomial terms. The multinomial theorem states that the number of monomials terms of a fixed order q in d variables is $\begin{pmatrix} q+d-1 \\ d-1 \end{pmatrix}$, which implies the number of monomial terms of degree at most p is

$$\dim \mathbb{P}^p(\Omega) = \sum_{q=0}^{p} \begin{pmatrix} q+d-1 \\ d-1 \end{pmatrix}. \tag{2}$$

Another useful polynomial space is $\mathbb{Q}^p(\Omega)$, the function space of polynomials where each term is a monomial of the form $x_1^{\alpha_1}\cdots x_d^{\alpha_d}$ for $x \in \Omega$, $\alpha \in \mathbb{N}^d$, and $\max_i \alpha_i \leqslant p$

$$\mathbb{Q}^p(\Omega) := \operatorname{span}\left\{ f : \Omega \to \mathbb{R} \,\middle|\, f(x) = x_1^{\alpha_1}\cdots x_d^{\alpha_d},\, \alpha \in \mathbb{N}^d,\, \max_i \alpha_i \leqslant p \right\}. \tag{3}$$

This definition of $\mathbb{Q}^p(\Omega)$ emphasizes that it is a finite-dimensional vector space. Since the collection of monomials $\left\{x_1^{\alpha_1}\cdots x_d^{\alpha_d} \mid \alpha \in \mathbb{N}^d,\, \max_i \alpha_i \leqslant p\right\}$ is a basis for $\mathbb{Q}^p(\Omega)$ (spans the whole space and all monomials are linearly independent), the dimension $\mathbb{Q}^p(\Omega)$ is equal to the number of monomial terms. Since each independent variable can have $p+1$ different exponents $(0, 1, \dots, p)$ and there are d variables, we have

$$\dim \mathbb{Q}^p(\Omega) = (p+1)^d. \tag{4}$$

In one dimension ($d = 1$) it is easy to verify that the space $\mathbb{P}^p(\Omega)$ and $\mathbb{Q}^p(\Omega)$ are equal and $\dim \mathbb{P}^p(\Omega) = \dim \mathbb{Q}^p(\Omega) = p + 1$.

Example: One variable

The dimension of the polynomial spaces in one variable are: $\dim \mathbb{P}^p = \dim \mathbb{Q}^p = p+1$

$$\mathbb{P}^0 = \mathbb{Q}^0 = \operatorname{span}\{1\}, \quad \mathbb{P}^1 = \mathbb{Q}^1 = \operatorname{span}\{1, x_1\}, \quad \mathbb{P}^2 = \mathbb{Q}^2 = \operatorname{span}\{1, x_1, x_1^2\} \tag{5}$$

Example: Two variables

The dimension of the polynomial spaces in two variables are: $\dim \mathbb{P}^p = (p+1)(p+2)/2$ and $\mathbb{Q}^p = (p+1)^2$

$$\begin{aligned} &\mathbb{P}^0 = \operatorname{span}\{1\}, && \mathbb{P}^1 = \operatorname{span}\{1, x_1, x_2\}, && \mathbb{P}^2 = \operatorname{span}\{1, x_1, x_2, x_1^2, x_1x_2, x_2^2\}, \\ &\mathbb{Q}^0 = \operatorname{span}\{1\}, && \mathbb{Q}^1 = \operatorname{span}\{1, x_1, x_2, x_1x_2\}, && \mathbb{Q}^2 = \operatorname{span}\{1, x_1, x_2, x_1x_2, x_1^2, x_2^2, x_1^2x_2, x_1x_2^2, x_1^2x_2^2\} \end{aligned} \tag{6}$$

2 Bases of Polynomial Spaces $\mathbb{P}^p$, $\mathbb{Q}^p$

Let $\mathcal{P}$ denote either the polynomial space $\mathbb{P}^p$ or $\mathbb{Q}^p$. From the previous section, we have that $\mathcal{P}$ is a finite-dimensional vector space with dimension

$$\dim \mathcal{P} = \begin{cases} \displaystyle\sum_{q=0}^{p} \binom{q+d-1}{d-1} & \text{if } \mathcal{P} = \mathbb{P}^p \\ (p+1)^d & \text{if } \mathcal{P} = \mathbb{Q}^p. \end{cases} \tag{7}$$

Recall that a basis of a vector space is a collection of linearly independent functions that span the space. Any basis of $\mathcal{P}$ must have exactly $\dim \mathcal{P}$ functions. As with all vector spaces, the choice of basis is non-unique; standard choices for a basis include monomials, orthogonal polynomials, and nodal polynomials. In infinite precision, all bases are equivalent since they are simply different representations of the same space. However, in finite precision, differences do arise due to conditioning issues (terms becoming nearly linearly dependent). In addition, the various basis choices have properties that make them more or less convenient in practice. In this course, we focus primarily on *nodal polynomial bases* and, to a lesser extent, monomial bases.

2.1 Monomial basis

A monomial basis of $\mathcal{P}$ is a collection of independent monomials with a specific property. It is the simplest and most intuitive basis of any polynomial vector space; in fact, the spaces $\mathbb{P}^p$ and $\mathbb{Q}^p$ were *defined* as the space spanned by specific monomial terms in Section 1. Despite the simplicity of monomial bases, they are not well-suited for a computational setting since, as the degree of the polynomial space increases, the monomial terms become *nearly* linear dependent (although still independent in infinite precision), which leads to an ill-conditioned system and computations will be dominated by round-off errors. However, they will be useful in *constructing* other bases of $\mathcal{P}$, as we will see in the next section.

2.2 Nodal polynomial basis

A nodal polynomial basis is a basis of $\mathcal{P}$, denoted $\{\psi_i(x)\}_{i=1}^{\dim \mathcal{P}}$, generated by a collection of N_v unique *nodes* $\{\hat{x}_i\}_{i=1}^{N_v}$ where $\hat{x}_i \in \Omega \subset \mathbb{R}^d$ for $i = 1, \dots, N_v$ and $\hat{x}_i \neq \hat{x}_j$ if $i \neq j$. To each node $\hat{x}_i$ we assign a corresponding basis function $\psi_i(x) \in \mathcal{P}$ and require

$$\psi_i(\hat{x}_j) = \delta_{ij} \tag{8}$$

for $i = 1, \dots, \dim \mathcal{P}$ and $j = 1, \dots, N_v$, which we will refer to as the *Lagrangian property*. Observe for each basis function $\psi_i(x)$, the Lagrangian property in (8) imposes N_v independent conditions. Since $\psi_i \in \mathcal{P}$, it is a linear combination of $\dim \mathcal{P}$ monomials and therefore *uniquely* determined by these N_v conditions provided $N_v = \dim \mathcal{P}$, i.e., it can be written as a linear system of $\dim \mathcal{P} = N_v$ equations in $\dim \mathcal{P} = N_v$ variables and will be invertible provided the nodes are unique. In the remainder of this section, we construct nodal bases for $\mathbb{P}^p$ and $\mathbb{Q}^p$. First we consider the special case of a single variable ($d = 1$) and then the special case of multiple variables where the nodes lie on an orthogonal grid in $\mathbb{R}^d$. We conclude with a general approach to construct nodal bases for $\mathbb{P}^p$ and $\mathbb{Q}^p$.

2.2.1 Construction of nodal basis: $\mathbb{P}^p = \mathbb{Q}^p$, one variable

Recall in the special case of one variable ($d = 1$), we have that $\mathcal{P} = \mathbb{P}^p = \mathbb{Q}^p$ and $\dim \mathcal{P} = p + 1$. In this special case, a nodal polynomial basis can be constructed by inspection. We begin with two examples.

Example: Nodal basis of $\mathbb{P}^1 = \mathbb{Q}^1$

Consider two distinct nodes in $\mathbb{R}$: $\{\hat{x}_1, \hat{x}_2\}$. The linear polynomials $\psi_1(x), \psi_2(x)$ that satisfy the Lagrangian property $\psi_1(\hat{x}_1) = 1$, $\psi_1(\hat{x}_2) = 0$, $\psi_2(\hat{x}_1) = 0$, $\psi_2(\hat{x}_2) = 1$ are

$$\psi_1(x) = \frac{x - \hat{x}_2}{\hat{x}_1 - \hat{x}_2}, \quad \psi_2(x) = \frac{x - \hat{x}_1}{\hat{x}_2 - \hat{x}_1}. \tag{9}$$

Since the Lagrangian property imposes two constraints on each basis function, each basis function is uniquely defined (a linear function is uniquely defined by two points it passes through). See Figure 1 for a plot of the basis functions.

Example: Nodal basis of $\mathbb{P}^2 = \mathbb{Q}^2$

Consider three distinct nodes in $\mathbb{R}$: $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$. The quadratic polynomials $\psi_1(x), \psi_2(x), \psi_3(x)$ that satisfy the Lagrangian property

$$\begin{aligned} &\psi_1(\hat{x}_1) = 1, \quad \psi_1(\hat{x}_2) = 0, \quad \psi_1(\hat{x}_3) = 0, \\ &\psi_2(\hat{x}_1) = 0, \quad \psi_2(\hat{x}_2) = 1, \quad \psi_2(\hat{x}_3) = 0, \\ &\psi_3(\hat{x}_1) = 0, \quad \psi_3(\hat{x}_2) = 0, \quad \psi_3(\hat{x}_3) = 1 \end{aligned} \tag{10}$$

are

$$\psi_1(x) = \frac{(x-\hat{x}_2)(x-\hat{x}_3)}{(\hat{x}_1-\hat{x}_2)(\hat{x}_1-\hat{x}_3)}, \quad \psi_2(x) = \frac{(x-\hat{x}_1)(x-\hat{x}_2)}{(\hat{x}_2-\hat{x}_1)(\hat{x}_2-\hat{x}_3)}, \quad \psi_3(x) = \frac{(x-\hat{x}_1)(x-\hat{x}_2)}{(\hat{x}_3-\hat{x}_1)(\hat{x}_3-\hat{x}_2)}. \tag{11}$$

Since the Lagrangian property imposes three constraints on each basis function, each basis function is uniquely defined (a quadratic function is uniquely defined by three points it passes through). See Figure 1 for a plot of the basis functions.

In both cases, we constructed the basis function $\psi_i(x)$ as a product of terms of the form $(x - \hat{x}_j)$ for $j = 1, \dots, i-1, i+1, \dots, p+1$, which is a polynomial of degree p and passes through zero at $\hat{x}_1, \dots, \hat{x}_{i-1}, \hat{x}_i, \dots, \hat{x}_{p+1}$. The only condition that remains to be satisfied is the normalization of the polynomial at $\hat{x}_i$, i.e., $\psi_i(\hat{x}_i) = 1$ (no sum over i), which comes from dividing the previous product by $(\hat{x}_i - \hat{x}_j)$ for $j = 1, \dots, i-1, i+1, \dots, p+1$. Therefore, we have constructed $\psi_i(x)$ as a polynomial of degree p with the Lagrangian property, i.e., passes through $p+1$ points (1 at $\hat{x}_i$ and 0 at all other nodes) and therefore is the *unique* polynomial of degree $p+1$ that accomplishes this. The general formula for the basis funtion $\psi_i(x)$ associated with node $\hat{x}_i$ is

$$\psi_i(x) = \prod_{\substack{j=1 \\ j \neq i}}^{p+1} \frac{x - \hat{x}_j}{\hat{x}_i - \hat{x}_j} \tag{12}$$

for $i = 1, \dots, p+1$. By substituting node $\hat{x}_k$ into (12) the Lagrangian property is easily verified. The polynomials in (12) are called the *Lagrange polynomials*. The Lagrange polynomials for $p = 1, 2, 3$ are shown in Figure 1.

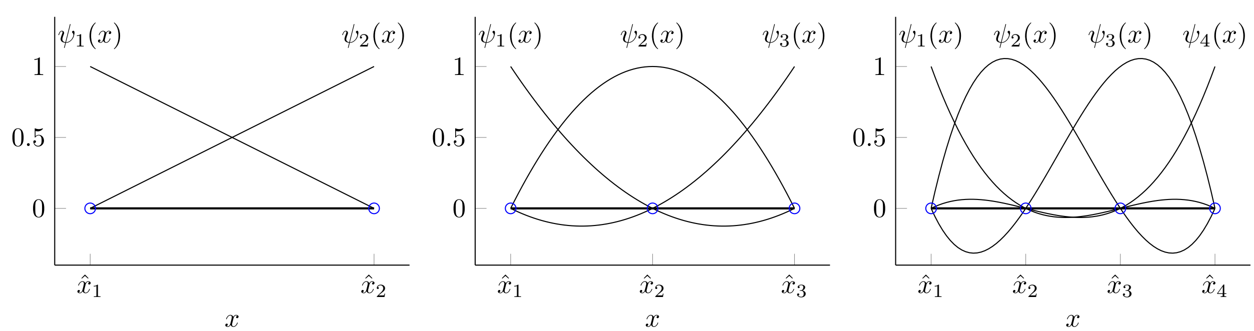

Figure 1: Lagrange polynoimals for $p = 1$ (*left*), $p = 2$ (*center*), $p = 3$ (*right*).

The derivative of the Lagrange polynomials comes from a careful application of the product rule of differentiating polynomials

$$\frac{d\psi_i}{dx}(x) = \sum_{k=1}^{p+1} \frac{1}{\hat{x}_i - \hat{x}_k} \prod_{\substack{j=1 \\ j \neq i,k}}^{p+1} \frac{x - \hat{x}_j}{\hat{x}_i - \hat{x}_j} \tag{13}$$

for $i = 1, \ldots, p+1$. The Lagrange polynomials possess the following *partition of unity* property for any $x \in \Omega$

$$\sum_{i=1}^{p+1} \psi_i(x) = 1, \quad \sum_{i=1}^{p+1} \frac{d\psi_i}{dx}(x) = 0. \tag{14}$$

The first property can be shown by summing the Lagrangian polynomials and the second property follows trivially by differentiating the first equation.

2.2.2 Construction of nodal basis: $\mathbb{Q}^p$, multiple variables, regular nodes

Next we consider construction of a basis for $\mathcal{P} = \mathbb{Q}^p$ in d variables ($d \geqslant 1$) in the special case where the generating nodes have a regular arrangement, i.e., lie on an orthogonal lattice. Recall $\dim \mathcal{P} = (p+1)^d$ in the special case where $\mathcal{P} = \mathbb{Q}^p$ and, for notational simplicity, we restrict our attention to the special case of $d = 2$; however, the construction will apply in any dimension $d \geqslant 1$. Introduce a set of points $\{\hat{s}_i\}_{i=1}^{p+1}$ in $\mathbb{R}$, i.e., for each $i = 1, \ldots, p+1$, $\hat{s}_i \in \mathbb{R}$, that will be used to generate a lattice of nodes $\{\hat{x}_{ij}\}_{i,j=1}^{p+1}$ as $\hat{x}_{ij} = (\hat{s}_i, \hat{s}_j)^T$ (Figure 2).

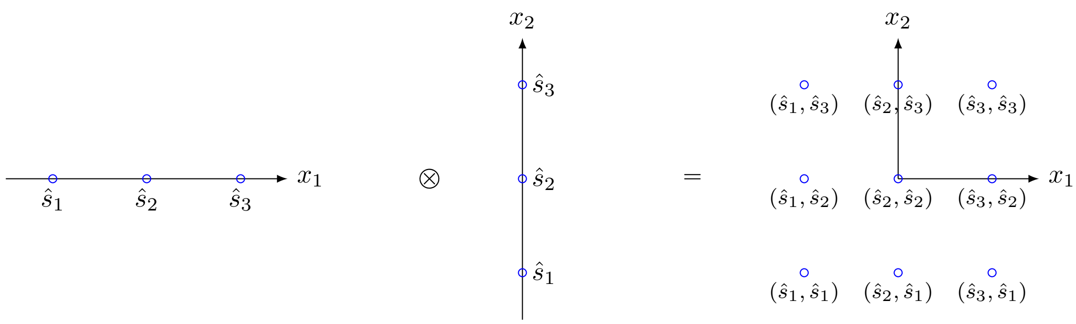

Figure 2: Nodal distribution in two dimensions as a tensor product of nodes in one dimension.

To each of these nodes, we associate a basis function $\chi_{ij}(x) \in \mathbb{Q}^p$, defined as

$$\chi_{ij}(x) = \varphi_i(x_1)\varphi_j(x_2), \tag{15}$$

where $\{\varphi_i\}_{i=1}^{p+1}$ are the Lagrange polynomials with generating nodes $\{\hat{s}_i\}_{i=1}^{p+1}$, i.e., $\varphi_i(\hat{s}_j) = \delta_{ij}$ for $i, j = 1, \ldots, p+1$. This choice of $\chi_{ij}(x)$ satisfies the Lagrangian property

$$\chi_{ij}(\hat{x}_{kl}) = \varphi_i(\hat{s}_k)\varphi_j(\hat{s}_l) = \delta_{ik}\delta_{sl}, \tag{16}$$

i.e., $\chi_{ij}(x)$ takes the value 1 at its corresponding node $\hat{x}_{ij}$ and zero at all other nodes. Furthermore, since the Lagrangian property imposes $(p+1)^d$ independent constraints on $\chi_{ij} \in \mathbb{Q}^p$ and $\dim \mathbb{Q}^p = (p+1)^d$, each basis function is unique.

The partial derivatives of the nodal basis functions are

$$\frac{\partial \chi_{ij}}{\partial x_1}(x) = \varphi_i'(x_1)\varphi_j(x_2), \qquad \frac{\partial \chi_{ij}}{\partial x_2}(x) = \varphi_i(x_1)\varphi_j(x_2)', \tag{17}$$

where $\varphi'(x)$ is the derivative of the univarite function $\varphi(x)$. The partition of unity property is inherited directly from the Lagrange polynomials in one variable

$$\sum_{i,j=1}^{p+1} \chi_{ij}(x) = \sum_{i,j=1}^{p+1} \varphi_i(x_1)\varphi_j(x_2) = \left(\sum_{i=1}^{p+1} \varphi_i(x_1)\right)\left(\sum_{j=1}^{p+1} \varphi_j(x_2)\right) = 1, \tag{18}$$

which immediately leads to the corresponding result for the partial derivatives

$$\sum_{i,j=1}^{p+1} \frac{\partial \chi_{ij}}{\partial x_1} = \sum_{i,j=1}^{p+1} \frac{\partial \chi_{ij}}{\partial x_2} = 0. \tag{19}$$

The nodal basis for $\mathbb{Q}^p$ has been introduced using multiple indices to conveniently leverage the tensor product structure of the nodal distribution. To be consistent with the remainder of the document, we convert this to a single (linear) index: define $\{\psi_i\}_{i=1}^{(p+1)^d}$, where

$$\psi_k(x) = \chi_{ij}(x), \quad k = i + (p+1)(j-1) \tag{20}$$

for $i = 1, \dots, p+1$ and $j = 1, \dots, p+1$.

Example: Nodal basis of $\mathbb{Q}^1$

Consider the one-dimensional nodes $\{\hat{s}_1, \hat{s}_2, \hat{s}_3\}$. The polynomials $\chi_{11}, \chi_{21}, \chi_{12}, \chi_{22} \in \mathbb{Q}^1$ that satisfy the Lagrangian property (16) are

$$\begin{aligned} \chi_{11}(x) &= \frac{(x_1 - \hat{s}_2)(x_2 - \hat{s}_2)}{(\hat{s}_1 - \hat{s}_2)(\hat{s}_1 - \hat{s}_2)}, \qquad & \chi_{21}(x) &= \frac{(x_1 - \hat{s}_1)(x_2 - \hat{s}_2)}{(\hat{s}_2 - \hat{s}_1)(\hat{s}_1 - \hat{s}_2)}, \\ \chi_{12}(x) &= \frac{(x_1 - \hat{s}_2)(x_1 - \hat{s}_1)}{(\hat{s}_1 - \hat{s}_2)(\hat{s}_2 - \hat{s}_1)}, \qquad & \chi_{22}(x) &= \frac{(x_1 - \hat{s}_1)(x_2 - \hat{s}_1)}{(\hat{s}_2 - \hat{s}_1)(\hat{s}_2 - \hat{s}_1)}, \end{aligned} \tag{21}$$

which can easily be verified by evaluating each basis function at the nodes

$$\hat{x}_{11} = (\hat{s}_1, \hat{s}_1)^T, \qquad \hat{x}_{21} = (\hat{s}_2, \hat{s}_1)^T, \qquad \hat{x}_{12} = (\hat{s}_1, \hat{s}_2)^T, \qquad \hat{x}_{22} = (\hat{s}_2, \hat{s}_2)^T. \tag{22}$$

To convert to a linear index we introduce $\{\psi_i\}_{i=1}^4$ as

$$\psi_1 = \chi_{11}, \qquad \psi_2 = \chi_{21}, \qquad \psi_3 = \chi_{12}, \qquad \psi_4 = \chi_{22}. \tag{23}$$

2.2.3 Construction of nodal basis: $\mathbb{P}^p$ and $\mathbb{Q}^p$, multiple variables

In the general case, a nodal basis of $\mathbb{P}^p$ or $\mathbb{Q}^p$ can be constructed using Vandermonde's method; see project notes for a description.

2.3 Generalized nodal polynomial basis

The choice to place constraints on only the *values* of the of the basis functions at each node in the definition of a nodal basis was made for the sake of simplicity; however, the concept of a nodal basis can be generalized to impose constraints on the value and first K derivatives of the basis functions. This concept is useful in the construction of basis functions for the finite element method applied to partial differential equations of order 4 or higher; however, it is beyond the scope of this class.

3 Polynomial approximation

Suppose we wish to approximate some function $f : \Omega \to \mathbb{R}$, where $\Omega \subset \mathbb{R}^d$ using polynomials. One way to approach this problem is to construct a basis $\{\psi_i(x)\}_{i=1}^{\dim \mathcal{P}}$ of a chosen polynomial space $\mathcal{P}$, e.g., $\mathbb{P}^p(\Omega)$ or $\mathbb{Q}^p(\Omega)$, and approximate f in that basis

$$f(x) \approx f_{\mathcal{P}}(x) := \sum_{i=1}^{\dim \mathcal{P}} \beta_i \psi_i(x), \tag{24}$$

where $p_m(x)$ is the polynomial approximation of f, and $\{\beta_i\}_{i=1}^m$ are the coefficients of the expansion yet to be determined. To determine the values of the coefficients, we must enforce that $f(x)$ and $f_{\mathcal{P}}(x)$ match in some sense, e.g., the moments of the functions match or their values and/or derivatives agree at some selected points. We focus on the later choice of requiring the function and our polynomial approximation agree at some selected points in Ω, which is commonly referred to as *polynomial interpolation.*

Let $\{\hat{x}_i\}_{i=1}^{\dim \mathcal{P}}$ be a set of unique points in Ω and let $\{\psi_i\}_{i=1}^{\dim \mathcal{P}}$ be the nodal basis generated by the nodes. Since we have chosen a nodal basis, the requirement that the function f and the polynomial interpolant match at the nodes simply reduces to

$$f(\hat{x}_i) = f_{\mathcal{P}}(\hat{x}_i) = \sum_{s=1}^{\dim \mathcal{P}} \beta_s \psi_s(\hat{x}_i) = \beta_i, \tag{25}$$

i.e., the coefficients must be taken as the function evaluated at the corresponding node.