Approximation Theory II
MATH46052|66052 *

David J. Silvester
School of Mathematics, University of Manchester

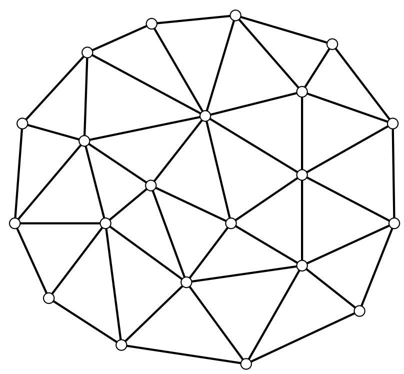
Version 1.0 | 5 February 2019

Contents

*These notes were written by Prashanth Nadukandi (`https://nadukandi.es`)

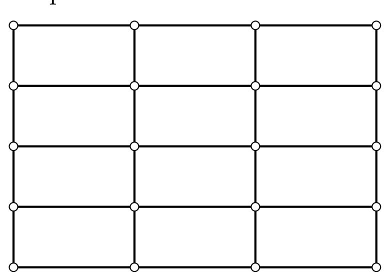
1 Piecewise approximation in $\mathbb{R}^2$

Two-dimensional piecewise polynomial approximation is the basis of the finite element method for solving partial differential equations.



The data points in two-dimensional space are members of a set of triples $\{(x_k, y_k, f_k)\}_{k=0}^{n_k}$, where (x_k, y_k) are the position coordinates of the kth point and f_k is the underlying function value at the kth point. Our objective is to construct a continuous approximation in two dimensions that interpolates these data points with a specified smoothness. In our case, the domain Ω will always be a polygonal region enclosing the set $\{(x_k, y_k)\}_{k=0}^{n_k}$ such that vertices on the boundary belong to the aforesaid set. There is no unique way of doing this. When the domain boundary $\partial\Omega$ is an arbitrary polygon, we will triangulate the interior region as shown in the adjacent picture. Given such a partition of Ω, we will construct the piecewise linear $C^0(\overline{\Omega})$ interpolant of the scattered data points.

When the boundary $\partial\Omega$ is rectangular and when the data points are placed in a structured arrangement of rows and columns, we will also consider a partition into several rectangular subdomains as shown in the picture below. The spacing between rows and between columns need not be uniform. Given such a grid of points we will construct the piecewise bilinear $C^0(\overline{\Omega})$ interpolant of the structured data. Additionally, if point values of the first derivatives and the cross derivative of the underlying function are also provided at the set of points, then we can construct the piecewise bicubic $C^1(\overline{\Omega})$ interpolant of the structured data points.



For structured data, when there are $m+1$ rows and $n+1$ columns then we have $(m+1)(n+1)$ points in two dimensions and $n_k = (m+1)(n+1) - 1$. The global index k has an ordered pair representation (i, j) where the index $i \in 0, 1, \ldots, m$ runs over the rows and the index $j \in 0, 1, \ldots, n$ runs over the columns. Further, we have the mapping $k \leftrightarrow (n+1)i + j$. Observe that the set $\{(x_i, y_j)\}_{i=0,j=0}^{i=m,j=n}$ is a Cartesian product of the one-dimensional sets $\{(x_i)\}_{i=0}^{m}$ and $\{(y_j)\}_{j=0}^{n}$. This directly leads one to the following definition.

(Tensor product spaces)
Let U and V be two linear function spaces. The tensor product of U and V is the function space spanned by the outer product of the set of basis functions of U and V.

1.1 Piecewise bilinear interpolation

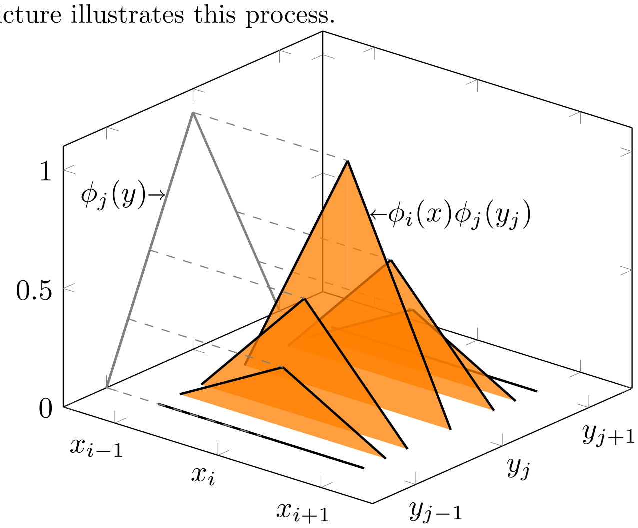
The space of piecewise bilinear functions defined on a rectangular grid is the simplest example of a tensor product space. For each point (x_i, y_j) on the grid we can identify a pair of point sets—one on the line parallel to the x-axis and the other on the line parallel to the y-axis passing though (x_i, y_j). Note that $\phi_i(x)$ is the ith global basis function for the piecewise linear function space defined on the former point set and so it takes the shape illustrated in part I of the notes. Likewise, $\phi_j(y)$ is the jth global basis function for the piecewise linear function space defined on the latter point set. In this case, the global basis function of (x_i, y_j) for the tensor product space is $\phi_i(x)\phi_j(y)$.

Let us denote by $f_{i,j}$ the given value of the underlying function at (x_i, y_j). Then the $C^0(\overline{\Omega})$ piecewise bilinear interpolant of the structured data is

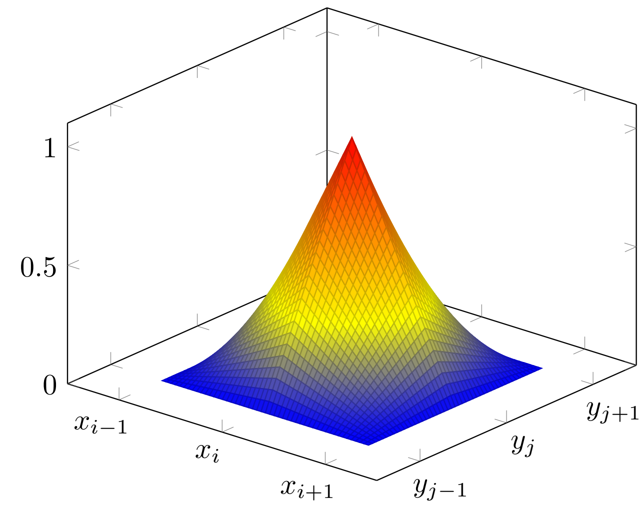
$$S_{1\otimes 1}(x,y) = \sum_i \sum_j f_{i,j}\phi_i(x)\phi_j(y). \tag{1.1}$$

Such simplicity!

One way to construct the surface of $\phi_i(x)\phi_j(y)$ is by sweeping the graph of $\phi_i(x)$ in the y-direction and while doing so multiplying its height by $\phi_j(y)$. The following picture illustrates this process.



The picture shows five snaphots of this construction, where the graph of $\phi_j(y)$ is displayed on the wall and the dashed lines indicate the height of $\phi_i(x)$ at each snapshot. It follows that the basis function $\phi_i(x)\phi_j(y)$ is zero outside the rectangular region whose corner points are (x_{i-1}, y_{i-1}) and (x_{i+1}, y_{i+1}). Further, as both $\phi_i(x)$ and $\phi_j(y)$ are continuous, the surface of $\phi_i(x)\phi_j(y)$ is continuous as shown in the following picture.



Observe that any cross section taken parallel to either the x-axis or the y-axis is piecewise linear. Hence the bilinear basis function $\phi_i(x)\phi_j(y)$ has a ruled surface. Ruled surfaces can have nonzero curvature which is evident in the surface plot of $\phi_i(x)\phi_j(y)$ shown.

1.2 Piecewise bicubic interpolation

Suppose that the function values, partial derivatives and cross derivatives at the nodes of any rectangular grid are given. Then, using the one-dimensional cubic Hermite basis functions illustrated in part I of the notes. we can construct the tensor product space in two dimensions whose basis functions span the piecewise bicubic $C^1(\overline{\Omega})$ function space.

Recall that in one dimension, there are two distinct global basis functions $\widehat{\phi}_i(x)$ and $\widetilde{\phi}_i(x)$ for the ith node. So for the tensor product space we get four global basis functions at every interior node of the grid:

$$\widehat{\phi}_i(x)\widehat{\phi}_j(y), \quad \widehat{\phi}_i(x)\widetilde{\phi}_j(y), \quad \widetilde{\phi}_i(x)\widehat{\phi}_j(y), \quad \widetilde{\phi}_i(x)\widetilde{\phi}_j(y).$$

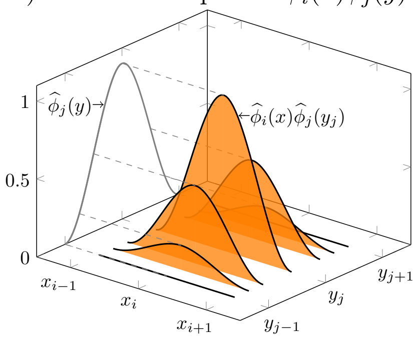
Next, let us define the notation,

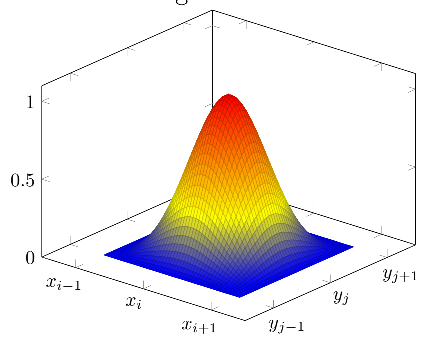
$$f^{\circ\circ}_{i,j} \coloneqq f(x_i, y_j),\ f^{\bullet\circ}_{i,j} \coloneqq \frac{\partial f}{\partial x}(x_i, y_j),\ f^{\circ\bullet}_{i,j} \coloneqq \frac{\partial f}{\partial y}(x_i, y_j),\ f^{\bullet\bullet}_{i,j} \coloneqq \frac{\partial^2 f}{\partial x \partial y}(x_i, y_j);$$

then the piecewise bicubic $C^1(\Omega)$ interpolant of structured data is given by

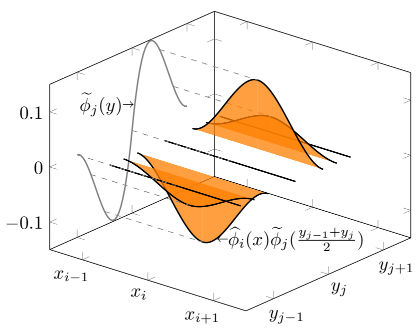
$$\begin{aligned} S_{2\otimes 2}(x,y) = {} & \sum_i \sum_j f^{\circ\circ}_{i,j}\widehat{\phi}_i(x)\widehat{\phi}_j(y) + f^{\bullet\circ}_{i,j}\widetilde{\phi}_i(x)\widehat{\phi}_j(y) \\ & + \sum_i \sum_j f^{\circ\bullet}_{i,j}\widehat{\phi}_i(x)\widetilde{\phi}_j(y) + f^{\bullet\bullet}_{i,j}\widetilde{\phi}_i(x)\widetilde{\phi}_j(y). \end{aligned} \tag{1.2}$$

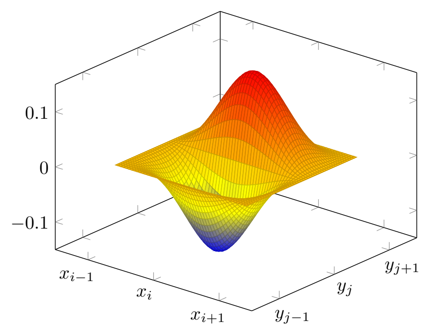
Exactly as above, we can construct the surface of $\widehat{\phi}_i(x)\widehat{\phi}_j(y)$ by sweeping the graph of $\widehat{\phi}_i(x)$ in the y-direction and while doing so multiplying its height by $\widehat{\phi}_j(y)$. Five snapshots taken during this process are shown in the picture (left). The surface plot of $\widehat{\phi}_i(x)\widehat{\phi}_j(y)$ is shown on the right.





To fix ideas and aid visualization, we show similar plots of the global basis function $\widehat{\phi}_i(x)\widetilde{\phi}_j(y)$:



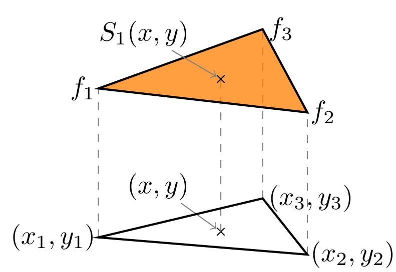


Observe that the two global basis functions shown in the two pictures above are continuous and are zero outside the patch of four rectangles that meet at (x_i, y_j). They are also $C^1(\overline{\Omega})$ continuous due to the smoothness properties inherited from the respective one-dimensional global basis functions. These properties also hold for the other two global basis functions given in (1.2).

Scattered data interpolation is discussed next.

1.3 Piecewise linear interpolation

Consider an arbitrary triangle in Ω with three vertices denoted by $\{(x_k, y_k)\}_{k=1}^3$. We denote by $S_1(x, y)$ the linear interpolant of the underlying function whose given values at the vertices are denoted by $\{f_k\}_{k=1}^3$. The adjacent picture illustrates the setting. We will derive the expression



for $S_1(x,y)$ and therein identify the
basis functions of linear interpolation on a triangle.

Observe in the previous picture that, if $S_1(x,y)$ is a linear interpolant, then the points $\{(x_k,y_k,f_k)\}_{k=1}^3$ and (x,y,S_1) lie on a plane. Using the fact that the scalar triple product of three vectors lying on a plane is zero gives the characterization

$$(x-x_1,y-y_1,S_1-f_1)\cdot\big((x_2-x_1,y_2-y_1,f_2-f_1)\times(x_3-x_1,y_3-y_1,f_3-f_1)\big)=0$$

which is equivalent to

$$\begin{vmatrix} x-x_1 & y-y_1 & S_1-f_1 \\ x_2-x_1 & y_2-y_1 & f_2-f_1 \\ x_3-x_1 & y_3-y_1 & f_3-f_1 \end{vmatrix}=0.$$

The above equation can be rewritten[1] as

$$\begin{vmatrix} x_1 & y_1 & f_1 & 1 \\ x & y & S_1 & 1 \\ x_2 & y_2 & f_2 & 1 \\ x_3 & y_3 & f_3 & 1 \end{vmatrix}=0,$$

and expanding the determinant along the third column we get

$$f_1\begin{vmatrix} x & y & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} - S_1\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} + f_2\begin{vmatrix} x_1 & y_1 & 1 \\ x & y & 1 \\ x_3 & y_3 & 1 \end{vmatrix} - f_3\begin{vmatrix} x_1 & y_1 & 1 \\ x & y & 1 \\ x_2 & y_2 & 1 \end{vmatrix}=0.$$

Rearranging this then gives the desired Lagrange interpolant:

$$S_1(x,y)=\underbrace{\frac{\begin{vmatrix} x & y & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}}{\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}}}_{L_1(x,y)} f_1 + \underbrace{\frac{\begin{vmatrix} x_1 & y_1 & 1 \\ x & y & 1 \\ x_3 & y_3 & 1 \end{vmatrix}}{\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}}}_{L_2(x,y)} f_2 + \underbrace{\frac{\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x & y & 1 \end{vmatrix}}{\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}}}_{L_3(x,y)} f_3. \tag{1.3}$$

Note that the terms $\{L_k(x,y)\}_{k=1}^3$ multiplying the given values $\{f_k\}_{k=1}^3$ are linear functions of x and y. Further, as S_1 is expressed as a linear combination of these terms, it follows that $\{L_k(x,y)\}_{k=1}^3$ are the canonical basis functions of linear interpolation on triangles. The values of $\{L_k(x,y)\}_{k=1}^3$ at any chosen (x,y) have a geometric interpretation as the ratios of two triangular areas. To see this, recall that

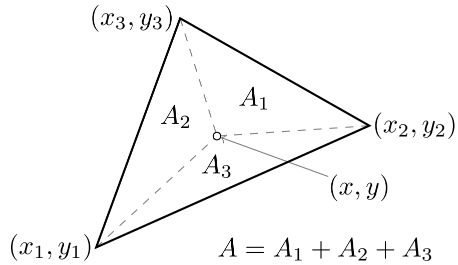
$$\begin{aligned}\text{abs}\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} &= \text{abs}\begin{vmatrix} x_2-x_1 & y_2-y_1 \\ x_3-x_1 & y_3-y_1 \end{vmatrix} \\ &= \text{abs}\big((x_2-x_1,y_2-y_1)\times(x_3-x_1,y_3-y_1)\big)\end{aligned}$$

[1] This can be seen by making the row operation $R_k \leftarrow R_k - R_1$ for $k=\{2,3,4\}$ and then expanding the determinant along the fourth column.

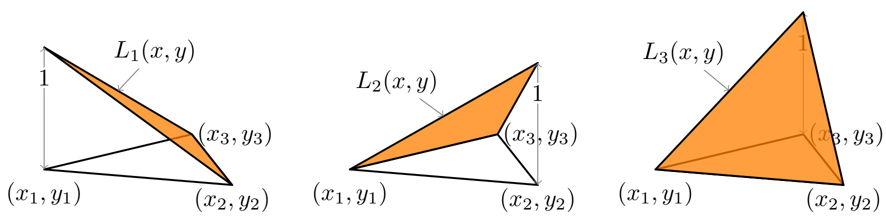
is twice the area of the triangle whose vertices are $\{(x_k, y_k)\}_{k=1}^{3}$.

Looking at the next picture, it can be readily seen that for any point (x, y) in the interior of a triangle of area A, we can construct three subtriangles of areas A_1, A_2 and A_3 such that

$$L_1(x,y) = \frac{A_1(x,y)}{A}, \quad L_2(x,y) = \frac{A_2(x,y)}{A}, \quad L_3(x,y) = \frac{A_3(x,y)}{A}. \tag{1.4}$$

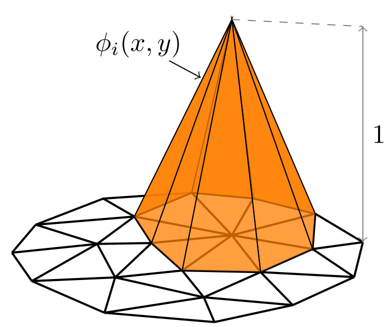


It follows that if (x, y) is an interior point then $0 \le L_k(x, y) \le 1$ for all k. Additionally, we have $L_1 + L_2 + L_3 = 1$. The following picture shows the three basis functions of linear interpolation on a triangle.



We note that for each $k \in \{1, 2, 3\}$, the basis function L_k has a unit value at the vertex (x_k, y_k) and is zero at the other two vertices. These results can be easily inferred from (1.4).

We now choose one vertex (let us call it the ith vertex) from the triangular mesh shown at the start of this section. To construct the ith global basis function of piecewise linear interpolation we first identify a patch of triangles that contain the vertex i. Then, on each triangle of the patch, the local version of $L_i(x, y)$ is constructed. Assembling the results gives the global basis function ϕ_i illustrated in the adjacent picture.

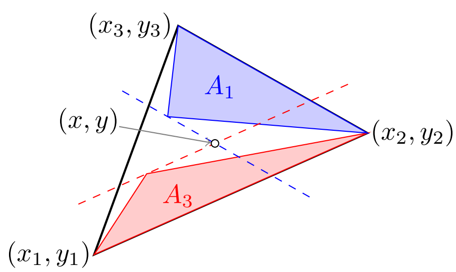


Observe that adjacent triangles on the patch share a common edge. the local basis functions from those adjacent triangles have identical graphs on the common edge—a straight line that is 1 at the ith vertex and is 0 at the other vertex. So, it follows that upon assembly the global basis function $\phi_i(x, y)$ is continuous and is zero outside the patch of eight triangles that meet at (x_i, y_j). As a result, the piecewise linear interpolant $S_1 \in C^0(\overline{\Omega})$ can be written in the standard form

$$S_1(x, y) = \sum_k f_k \phi_k(x, y). \tag{1.5}$$

The basis functions of linear interpolation on a triangle evaluated at any given point can be associated with the *barycentric coordinates* of that point in the triangle. We give a precise definition next.

(Barycentric coordinates)
The word barycentric has its origin in the Greek word *barus* which means heavy. If a triangle is considered to be of uniform density then the weight of any piece is proportional to its area. A point inside[2] a triangle can also be described using the areas A_1, A_2 and A_3 subtended by the edges at the point as the new coordinates. The use of three coordinates to describe a two-dimensional space involves redundancy only if they are all independent. Recall that $A_1 + A_2 + A_3 = A$. The following picture illustrates the fact that there is a bijective mapping $(x, y) \leftrightarrow (A_1, A_2, A_3)$.



The normalized coordinates $(A_1/A, A_2/A, A_3/A)$ are called the barycentric coordinates of a point on a triangle. Observe that the equation $A_1 = 0$ gives a simple expression for the equation of the line joining the vertices (x_2, y_2) and (x_3, y_3). Suppose we are asked to find the point described by the triple (A_1, A_2, A_3). The edge connecting (x_2, y_2) and (x_3, y_3) subtends an area A_1 (top shaded region) at all points on the dashed line parallel to that edge. The edge connecting (x_1, y_1) and (x_2, y_2) subtends an area A_3 (lower shaded region) at all points on the dashed line parallel to that edge. As the dashed lines are not parallel, there is one and only one point of intersection. Further, if $A_1 > 0$ and $A_3 > 0$, then the point of intersection lies strictly in the interior of the triangle.

The use of barycentric coordinates means that extending linear approximation to quadratic approximation on a triangle is a straightforward task.

[2] To describe points outside the triangle we need to consider negative areas; that is, an area with an orientation.

1.4 Piecewise quadratic interpolation

Any quadratic surface in two-dimensional space can be expressed as a linear combination of the quadratic monomial basis $1, x, y, xy, x^2, y^2$, so we need to determine six coefficients to uniquely define a quadratic surface. It follows that for quadratic interpolation on a triangle we need to specify interpolation conditions at six distinct points $\{(x_k, y_k)\}_{k=1}^{6}$ on the triangle.

If the triangular region is an element of a triangular mesh covering a domain, then the placement of the six points on each triangle needs to be done in such way that the global interpolating function is continuous. The continuity of a quadratic interpolating function across the common edge of adjacent triangles will be guaranteed if we have three distinct interpolation points on the edge. Thus, given a total of six distinct interpolation points per triangle, the only way we can place three points on each edge of the triangle is to position three of them at the vertices. The placement of the remaining three points, one on each edge, is motivated by the ease of construction of the quadratic Lagrange basis. The choice that results in the least bookkeeping is to position them at the midpoints of the edges.

We now denote by $\{N_k(x, y)\}_{k=1}^{6}$ the quadratic Lagrange basis functions that satisfy $N_i(x_j, y_j) = \delta_{ij}$. Since the monomial basis $\{1, x, y, xy, x^2, y^2\}$ is not a suitable Lagrange basis, we will build the quadratic basis set using the linear Lagrange basis functions (or, equivalently, the barycentric coordinates) $\{L_k(x, y)\}_{k=1}^{3}$ defined in the previous section.

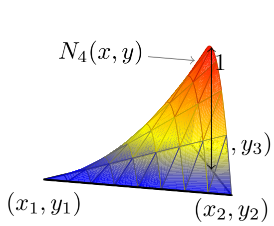
To this end, observe that span$\{L_1, L_2, L_3\}$ = span$\{1, x, y\}$ and since $L_i L_j$ is quadratic for all $i, j \in \{1, 2, 3\}$ we deduce that

$$\text{span}\,\{1, x, y, xy, x^2, y^2\} = \text{span}\,\{L_1, L_2, L_3, {L_1}^2, {L_2}^2, {L_3}^2, L_2L_3, L_3L_1, L_1L_2\}.$$

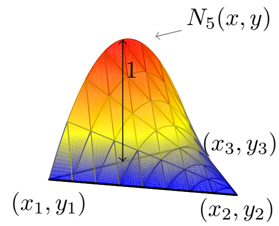
The functions on the right-hand-side of the above expression are not linearly independent. Specifically, from the definition of the barycentric coordinates we have $L_1 + L_2 + L_3 = 1$, so we can write $L_1 = {L_1}^2 + L_1L_2 + L_3L_1$, $L_2 = {L_2}^2 + L_2L_3 + L_1L_2$ and $L_3 = {L_3}^2 + L_3L_1 + L_2L_3$. This means that

$$\begin{aligned}\text{span}\,\{1, x, y, xy, x^2, y^2\} &= \text{span}\,\{L_1, L_2, L_3, L_2L_3, L_3L_1, L_1L_2\}\\ &= \text{span}\,\{{L_1}^2, {L_2}^2, {L_3}^2, L_2L_3, L_3L_1, L_1L_2\}.\end{aligned}$$

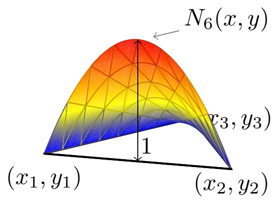
Observe that L_iL_j is zero valued at the three vertices and on all edges other than the one connecting the vertices i and j. On the latter edge, L_iL_j is nonnegative and has a parabolic profile. Recall that a parabola attains its extremum midway between the zeros of the parabola. Therefore, midway between the vertices i and j, L_iL_j attains a maximum value of $1/4$. Hence, after scaling (multiplying by 4) L_2L_3, L_3L_1 and L_1L_2 are representations of the quadratic Lagrange basis functions N_4, N_5 and N_6, respectively. They are pictured below.



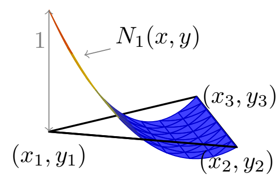
$\boxed{N_4 = 4L_2L_3}$



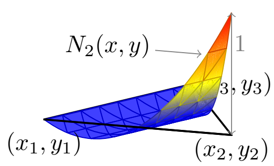
$\boxed{N_5 = 4L_3L_1}$



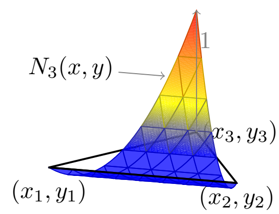
$\boxed{N_6 = 4L_1L_2}$



$\boxed{N_1 = 2{L_1}^2 - L_1}$



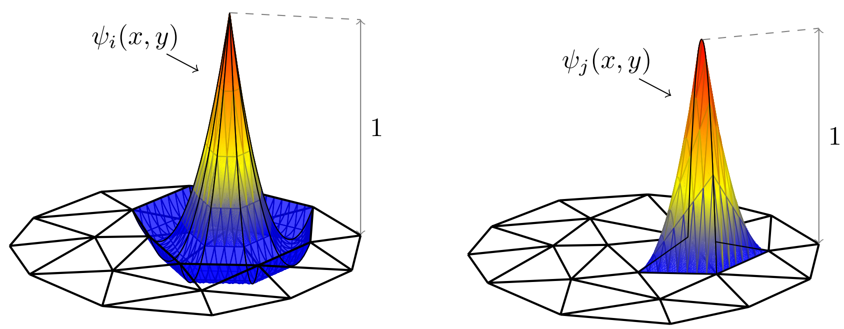
$\boxed{N_2 = 2{L_2}^2 - L_2}$



$\boxed{N_3 = 2{L_3}^2 - L_3}$

Unfortunately, ${L_1}^2$, ${L_2}^2$ and ${L_3}^2$ are not suitable candidates for the quadratic Lagrange basis function N_1, N_2 and N_3. This is because ${L_i}^2$ is not zero valued at the midpoint of the edges. An important geometric property of L_i is that its contours are straight lines parallel to the edge opposite the vertex i. Another geometric property of any triangle is that the line joining the midpoints of any two edges is parallel to the third edge. Therefore, the linear combination $sL_i + (1-s){L_i}^2$ will take value 1 on vertex i, value 0 on the opposite edge and constant values on lines parallel to the edge opposite to vertex i. Solving for s such that $sL_i + (1-s){L_i}^2$ is zero valued at the midpoints of the edges, we obtain $s = -1$. Thus, $2{L_1}^2 - L_1$, $2{L_2}^2 - L_2$ and $2{L_3}^2 - L_3$ are representations of the quadratic Lagrange basis functions N_1, N_2 and N_3, respectively. One feature that is clearly visible from the plots is that the edge functions N_4, N_5 and N_6 are positive everywhere in the interior of the triangle (as are the underlying linear interpolation functions). The three vertex functions N_1, N_2, N_3, in contrast, each have a line of points in the interior of the triangle where the function is zero.

We observed in the previous section that the global basis functions associated with piecewise linear interpolation on a triangular mesh are uniquely associated with the *vertices* of the triangulation. This association is lost when using piecewise quadratic interpolation. Instead, we have a set of global basis functions corresponding to the vertices and a second set of basis functions corresponding to the edges of the mesh. The assembly of the local basis functions to get global basis functions is unchanged: pick a vertex i (or an edge j), identify a patch of triangles that contain that vertex (or adjoin that edge) and then, for each triangle within the patch, construct the corresponding basis function $\psi_i(x, y)$ (or $\psi_j(x, y)$). The global basis functions generated in this manner are illustrated in the picture below.



We conclude this section of the notes by reviewing the smoothness of piecewise polynomials in two dimensions.

(Regularity of piecewise polynomials)
Given any finite-dimensional function space V_h with a basis set consisting of piecewise polynomials, the following regularity results hold:

$$V_h \subset C^0(\overline{\Omega}) \iff V_h \in H^1(\Omega), \tag{1.6a}$$

$$V_h \subset C^1(\overline{\Omega}) \iff V_h \in H^2(\Omega). \tag{1.6b}$$

The equivalence (1.6a) follows from the fact that if a function v is a polynomial of x and y in each triangle or rectangle of a subdivision of $\overline{\Omega}$ and is continuous across all the interior edges of the subdivision, then the first derivative components of ∇v exist and are piecewise continuous, so that $v \in H^1(\Omega)$. On the other hand, if v is not continuous across one of the interior edges then one or more derivative components in ∇v do not exist as functions in $L^2(\Omega)$ and thus $v \notin H^1(\Omega)$.