

Where Do We Stand on Meshfree Approximation Schemes?

In a previous posting, Timon provided a nice overview of meshfree methods—starting from SPH and leading up to some of the key developments over the past decade (diffuse element method, element-free Galerkin, reproducing kernel particle method/RKPM). Rather than present details on any particular method per se, here I focus on the most common approximations that are used in meshfree methods. In doing so, my goal is to bring out the commonalities, distinctions, and some recent perspectives and improved understanding that has come about in the realm of data approximation and its ties to meshfree methods. Where appropriate, I will try to point out how the properties of the approximant lead to positive or negative consequences when used within a Galerkin method. The important issues of imposing essential boundary conditions and numerical integration in Galerkin meshfree methods are also discussed. In the interest of space, equations are inlined and no figures are included. Links to cited references (journal articles, web resources, or author's web page) are provided; the full citation of the references is available [here](#). The reader will notice that the title of this post is an adaptation of Jaynes's (1979) article—[Where Do We Stand on Maximum Entropy?](#)

Given a set of nodes $\{\mathbf{x}_a\}$ ($a = 1$ to n) in \mathbf{R}^d , we construct an approximation for a scalar-valued function $u(\mathbf{x})$ in the form: $u^h(\mathbf{x}) = \varphi_a(\mathbf{x})u_a$ (Einstein summation convention implied). Most, if not all meshfree methods are based on some variant of either [radial basis functions \(RBFs\)](#), [moving least squares \(MLS\) approximants](#) (in computational geometry, [Levin's \(1998\) MLS approximant](#) is adopted), or natural neighbor-based (nn) interpolants such as [Sibson coordinates](#) and [Laplace interpolant](#). Recently, [maximum-entropy approximants](#) have also come to the forefront. The mathematical analysis of meshfree approximants has been carried out by [Babuska et al. \(2003\)](#). In meshfree methods, the construction of the nodal basis functions $\varphi_a(\mathbf{x})$ is independent of the background element structure (unlike finite elements), and different approaches are used to construct a linearly independent set of basis functions. In this sense, these approximations are referred to as *meshfree*. A brief description of the above schemes follows.

Radial Basis Functions

In the radial basis function approximation, φ_a is constructed from translates of a fixed radial function φ with centers at \mathbf{x}_a . If global polynomial reproducibility is desired, a polynomial term is added to the approximation, which engenders an additional side condition. For certain choices of $\varphi(\cdot)$, for example, Gaussian radial basis function $[\exp(-r^2/c^2)]$, multiquadrics $[(r^2 + c^2)^{1/2}]$, or thin-plate splines, the matrix $\mathbf{K}_{ab} = \varphi_a(\|\mathbf{x}_b - \mathbf{x}_a\|)$ is positive-definite and invertible, and hence the data interpolation problem has a unique solution for the coefficients u_a . Note that u^h interpolates, but the radial functions φ_a do not satisfy the Kronecker-delta property, $\varphi_a(\mathbf{x}_b) = \delta_{ab}$. In approximation theory, basis functions with the property $\varphi_a(\mathbf{x}_b) = \delta_{ab}$ are known as a cardinal basis. For a cardinal basis set, we immediately see that the basis functions are linearly independent. When a cardinal basis is used, the coefficients u_a are more commonly referred to as nodal values (finite element terminology). The use of RBFs in collocation-based meshfree methods was initiated by [Kansa \(1990\)](#), and collocation methods that are based on global (full matrix with exponential convergence) as well as compactly-supported RBFs are an active area of current research.

Moving Least Squares Approximants

In the standard least squares approach, given a polynomial basis with m terms (a quadratic basis in one dimension is $\mathbf{p}(\mathbf{x}) = \{1, x, x^2\}^T$), the best fit to nodal data u_a is sought. To this end, we let $u^h(\mathbf{x}) = \mathbf{p}^T \mathbf{a}$, where the constant parameter vector \mathbf{a} is found such that the error vector $\mathbf{P}^T \mathbf{a} - \mathbf{u}$ is minimized. Here, \mathbf{P} is a constant $m \times n$ matrix with the a -th column consisting of $\mathbf{p}(\mathbf{x}_a)$. As the objective (criterion) function, the square of the L_2 norm of the error is chosen: $I(\mathbf{a}) = 1/2(\mathbf{P}^T \mathbf{a} - \mathbf{u})^T (\mathbf{P}^T \mathbf{a} - \mathbf{u})$. This leads to a quadratic minimization problem, and hence a linear system of equations (normal equations) is obtained for the unknown vector \mathbf{a} .

In the moving least squares approximation, a local weighted least squares fit at each point \mathbf{x} is carried out. A non-negative compactly-supported weight function (derived from a Gaussian or polynomial/spline function) is associated

with each node: $w_a(\mathbf{x}) = w(\|\mathbf{x} - \mathbf{x}_a\|/d_a)$, where d_a is the radius of support (circular or tensor-product supports are typically used) of the nodal weight function. Instead of the standard least squares objective, a weighted quadratic least squares minimization problem is solved to determine $\mathbf{a}(\mathbf{x})$ (parameters are now functions of \mathbf{x}): $\mathbf{I}(\mathbf{a}) = 1/2(\mathbf{P}^T \mathbf{a} - \mathbf{u})^T \mathbf{W} (\mathbf{P}^T \mathbf{a} - \mathbf{u})$. Here, \mathbf{W} is a $n \times n$ matrix with non-zero entries $w_a(\mathbf{x})$ only on the diagonal of the a -th row. On carrying out the minimization, the MLS basis function vector is: $\boldsymbol{\phi}(\mathbf{x}) = \mathbf{B}^T(\mathbf{x})\mathbf{A}^{-1}(\mathbf{x})\mathbf{p}(\mathbf{x}) = \mathbf{B}^T(\mathbf{x})\boldsymbol{\alpha}(\mathbf{x})$, with $\mathbf{A}(\mathbf{x})\boldsymbol{\alpha}(\mathbf{x}) = \mathbf{p}(\mathbf{x})$, and $\mathbf{A}(\mathbf{x}) = \mathbf{P}\mathbf{W}(\mathbf{x})\mathbf{P}^T$ (moment matrix) and $\mathbf{B}(\mathbf{x}) = \mathbf{P}\mathbf{W}(\mathbf{x})$. The intermediate steps in the derivation, computer implementation of MLS basis functions, and reviews on its applications for partial differential equations (PDEs) can be found in the literature (see [Belytschko et al. \(1996\)](#), [Dolbow and Belytschko \(1999\)](#), and [Fries and Mathies \(2004\)](#) for details).

Two particular attractions of the MLS approach are: first, the approximation u^h reproduces all functions that are contained in the basis vector \mathbf{p} , which can include polynomials as well as non-polynomial functions (this has been used as a means for intrinsic enrichment, for e.g., by incorporating crack-tip functions within the basis vector); and secondly, if the weight function is C^k and the basis vector \mathbf{p} is smooth, then ϕ_a are also C^k , which is pertinent for higher-order gradient continua, phase transformations, and thin-plate and thin-shell analyses that place C^1 continuity requirements on the trial and test approximations. Construction of C^1 finite element bases on arbitrary meshes is in general a non-trivial task; use of subdivision finite elements ([Cirak et al., 1999](#)) is a promising alternative. The positive attributes of MLS are counter-balanced by the fact that nodal interpolation is lost, and furthermore on the boundary of the domain interior nodal basis functions have in general a non-zero contribution. So, in a standard Galerkin variational formulation, the condition that MLS test functions must vanish on an essential boundary is not met. Hence, modifications in the test function or in the variational form are required to impose essential boundary conditions. The weight function and its support size (must be above a lower bound to ensure that \mathbf{A} is invertible for all \mathbf{x} in the domain) are free parameters in an analysis—this parallels the choice of the 'shape parameter' c in RBF methods for the solution of partial differential equations (see [Wertz et al. \(2006\)](#)).

In lieu of what is to follow, we also mention an alternative formulation of MLS. The unconstrained minimization problem that we posed earlier for MLS can be recast in the so-called primal-dual framework. The vector $\boldsymbol{\alpha}(\mathbf{x})$ is the solution of the primal problem (P): $\max_{\boldsymbol{\alpha}} -M(\boldsymbol{\alpha}) = \min_{\boldsymbol{\alpha}} M(\boldsymbol{\alpha})$, with $M(\boldsymbol{\alpha}) = 1/2\boldsymbol{\alpha}^T \mathbf{A}\boldsymbol{\alpha} - \boldsymbol{\alpha}^T \mathbf{p}$ (pardon the abuse of notation), and the dual problem (D) is: $\min_{\boldsymbol{\phi}} D(\boldsymbol{\phi}) = 1/2\boldsymbol{\phi}^T \mathbf{W}^{-1} \boldsymbol{\phi}$ subject to the under-determined linear constraints $\mathbf{P}\boldsymbol{\phi} = \mathbf{p}$. The variables $\boldsymbol{\phi}$ (basis function vector) and $\boldsymbol{\alpha}$ (Lagrange multiplier vector) are related to each other via duality. The curious reader can write out the Lagrangian of the dual problem, set its first variation to zero and back-substitute the basis function vector in the Lagrangian functional to verify that the primal problem is obtained. From the dual problem (D), we clearly see that the reproducing conditions appear as equality constraints in the MLS approximation. Of course, the reproducibility of the basis vector $\mathbf{p}(\mathbf{x})$ is easily verified from the previous derivation: $\mathbf{P}\boldsymbol{\phi} = \mathbf{P}\mathbf{B}^T \mathbf{A}^{-1} \mathbf{p} = \mathbf{P}\mathbf{W}\mathbf{P}^T \mathbf{A}^{-1} \mathbf{p} = \mathbf{A}\mathbf{A}^{-1} \mathbf{p} = \mathbf{p}$. On a related note, if $\mathbf{W} = \mathbf{I}$ (identity matrix), then the minimum norm approximant ([Morse-Penrose or pseudo- or generalized-inverse](#), $\boldsymbol{\phi} = \mathbf{P}^+ \mathbf{p}$) is obtained. The MLS approximation can be viewed as a weighted minimum norm approximant, or equivalently the minimum Euclidean norm of the transformed vector $\mathbf{W}^{-1/2} \boldsymbol{\phi}$.

Natural Neighbor-Based Interpolants

For a set of nodes in \mathbf{R}^d , the Delaunay and Voronoi tessellation are dual geometric structures. Classical finite element bases are constructed on the Delaunay triangulation. On using the Voronoi diagram, [Sibson \(1980\)](#) introduced the concept of *natural neighbors* and natural neighbor (Sibson) interpolation. The Delaunay triangulation satisfies the empty circumcircle criterion (besides the vertex nodes of triangle T , no other nodes are located within the circumcircle of T). This property is used to define the *natural neighbors* of a point \mathbf{x} that is inserted within the convex hull of the nodal set. If \mathbf{x} lies within the circumcircle of a triangle T , then the vertex nodes of T are *natural neighbors* of \mathbf{x} . Let \mathbf{x} have n natural neighbors. Defining the area of overlap of the original Voronoi cell of node a with the Voronoi cell of point \mathbf{x} as $A_a(\mathbf{x})$ and the area of the Voronoi cell of point \mathbf{x} as $A(\mathbf{x})$, $\phi_a(\mathbf{x}) = A_a(\mathbf{x})/A(\mathbf{x})$, and the basis functions sum to unity by construction. A different natural neighbor interpolant was proposed by [Christ et al. \(1982\)](#), which was re-discovered in [applied mathematics](#) and [computational geometry](#). This interpolant (coined as Laplace

since it is a discrete solution to the Laplace equation) is constructed by using measures that are solely based on the Voronoi cell associated with \mathbf{x} . These interpolants are also linearly precise, and hence they are suitable for use within a Galerkin implementation for second-order PDEs. The appealing aspect of nn-interpolation is that it is well-defined and robust for very irregular distribution of nodes since the Voronoi diagram (and ergo natural neighbors) for a nodal set is unique. This is unlike the Delaunay triangulation, which is non-unique (four co-circular nodes in two dimensions leads to two possible triangulations and hence two different interpolants—data-dependent triangulation is well-known). The basis function supports automatically adapt (anisotropic supports) with changes in the nodal distribution, and hence no user-defined parameters are required to define nodal basis function supports. Further details on the construction of nn-interpolants are available [here](#). [Braun and Sambridge \(1995\)](#) introduced the use of the Sibson interpolant in a Galerkin method (*natural element method*), and many new and emerging applications of the method can be found [here](#).

Natural neighbor interpolation schemes share many common properties with the Delaunay finite element interpolant. They are linearly precise, strictly non-negative, and on convex domains they are piece-wise linear on the boundary. These permit the imposition of essential boundary conditions as in finite element methods. [Cueto et al. \(2000\)](#) combined Sibson interpolation with the concept of [\$\alpha\$ -shapes](#) to describe a domain discretized by a cloud of nodes and to track its evolution in large deformation analysis. The Sibson interpolant is $C^1 \setminus \mathbf{x}_a$ (derivatives are discontinuous at the nodes). Unlike MLS approximations, the development of higher-order continuous nn-interpolants is not straightforward. In this direction, [Farin \(1990\)](#) proposed a C^1 Sibson interpolant using the Bernstein-Bézier representation, and higher-order generalizations of nn-interpolants have also appeared (see [Hiyoshi and Sugihara \(2004\)](#)). An interesting advance due to [Boissonnat and Flötotto \(2004\)](#) is the extension of the Sibson interpolant to smooth approximations on a surface (($d-1$)-manifold in \mathbf{R}^d). An implementation of natural neighbor interpolation is available in the [Computational Geometry Algorithms Library \(CGAL\)](#).

Maximum-Entropy Approximants

In tracing the roots of data approximation, a common theme that emerges is that many approximants have a variational basis and are posed via an unconstrained or constrained optimization formulation. Cubic splines and thin-plate splines are prime examples, with MLS, RBFs, Laplace, discrete harmonic weight (see [Pinkall and Polthier \(1993\)](#)), and Kriging being a few notables that are linked to meshfree approximations. The reproducing conditions, $\mathbf{P}\boldsymbol{\phi} = \mathbf{p}$, have been the guiding principle behind the developments in meshfree (notably, RKPM of [Liu](#) and co-workers) and [partition of unity methods](#). In the RKPM, the basis function vector of the form $\boldsymbol{\phi}(\mathbf{x}) = \mathbf{W}\mathbf{P}^T(\mathbf{x})\boldsymbol{\alpha}(\mathbf{x})$ is considered; in the literature, often an additional multiplicative term (nodal volume) is included in the basis function definition. If the same nodal volume is assigned to each node, this approximation is identical to MLS. In general, the reproducing conditions can be seen as constraints, with the choice of the objective function being left open. In MLS, as was indicated earlier, a particular choice of the objective function was made. On imposing the requirement of linear precision, the problem is ill-posed in d dimensions if $n > d+1$. This is so since there are only $d+1$ equality constraints. As a means for regularization, an objective functional that is *least-biased* is desired. The [principle of maximum entropy](#) is a suitable candidate—initially used to demonstrate that Gibbs-Boltzmann statistics can be derived through inference and information theory, and in years thereafter has been successfully applied in many areas of pure and applied sciences where rationale inductive inference ([Bayesian theory of probability](#)) is required. In the presence of testable information (constraints) and when faced with epistemic (ignorance) uncertainty, the maximum entropy (MAXENT) formulation using the Shannon entropy functional ([Shannon \(1948\)](#), [Jaynes \(1957\)](#)) provides the least-biased statistical inference solution for the assignment of probabilities—[Wallis's combinatorial derivation](#) as well as the [maximum entropy concentration theorem](#) provide justification.

The Shannon entropy of a discrete probability distribution is: $H(\boldsymbol{\phi}) = -\sum_a \phi_a \ln \phi_a$. Historically, discrete probability measures have been seen as *weights* and hence their association with the construction of non-negative basis functions is natural. This led to the use of the maximum-entropy formalism to construct non-negative basis functions ([S, 2004](#), [Arroyo and Ortiz \[AO\], 2006](#)). These developments share common elements with the work of [Gupta \(2003\)](#) in supervised learning. In [S \(2004\)](#), the Shannon entropy is used within the maximum entropy variational principle to construct basis functions on polygonal domains, whereas in [AO \(2006\)](#), a modified entropy functional is adopted to construct local MAXENT approximation schemes for meshfree methods. The latter researchers noted its links to [convex analysis](#), and coined such approximants with the non-negative constraint, $\phi_a \geq 0$, as convex approximation schemes.

Natural neighbor-based interpolants as well as barycentric constructions on convex polygons are convex approximation schemes. The Delaunay interpolant is also the solution of an optimization problem, which was shown by [Rajan \(1991\)](#). The modified entropy functional is a linear combination (in the sense of Pareto optimality) of Rajan's functional and the Shannon entropy functional, and the solution of the variational problem provides a smooth transition from Delaunay interpolation as a limiting case at one end to global MAXENT approximation at the other end of the spectrum. Geometry has a lot to offer in computations, and once again, it is pleasing to see yet another connection emerge between geometry-and-approximation. Non-negative basis functions have many positive attributes (variation diminishing, convex hull property, positive-definite mass matrices, optimal conditioning), and their merits in computational mechanics has been recently demonstrated by [Hughes et al. \(2005\)](#) who used NURBS basis functions in isogeometric analysis.

A general prescription for locally- and globally-supported convex approximation schemes can be derived using the [Kullback-Leibler \(KL\) distance or directed divergence](#). This was introduced in [S \(2005\)](#) and is further elaborated in a forthcoming article. It was recognized (see [Jaynes \(2003\)](#)) that for the differential (continuous) entropy to be invariant under a transformation it must be of the form $-\phi \ln \phi/m dx$, which in the discrete case is: $H(\boldsymbol{\phi}, \mathbf{m}) = -\sum_a \phi_a \ln \phi_a / m_a$, where \mathbf{m} is a known prior distribution (*weight function*) for $\boldsymbol{\phi}$. The KL-distance, which is the negative of H , is non-negative (established using [Jensen's inequality](#)), and minimization of the relative entropy is the corresponding variational principle. We determine the non-negative basis functions, $\phi_a \geq 0$, by maximizing H , subject to the $d+1$ linear constraints, $\mathbf{P}\boldsymbol{\phi} = \mathbf{p}$. This is the *primal problem* for entropy maximization, which has a unique solution for any point \mathbf{x} within the convex hull of the nodal set. Outside the convex hull, the equality constraints and the non-negative restriction on the basis functions constitute an infeasible constraint set. To see this fact through a simple example, consider one-dimensional approximation with n nodes located in $[0,1]$ and let $x = -\delta$, where δ is positive. The first-order reproducing condition is: $\sum_a \phi_a x_a = -\delta$, and since all x_a are non-negative and $\delta > 0$, there does not exist any non-negative basis function vector $\boldsymbol{\phi}$ that can satisfy this constraint. The proof for the case when $x > 1$ proceeds along similar lines. The prior is a weight function that is chosen *a priori* (e.g., globally- or compactly-supported radial basis functions, weights used in MLS, [R-functions](#), etc.), and the above formulation provides a correction on the prior so that the basis functions satisfy the reproducing conditions. If a Gaussian radial basis function is used as a prior, then the modified entropy functional considered in [AO \(2006\)](#) is recovered.

On using the method of Lagrange multipliers, the MAXENT basis functions are obtained (exponential form): $\phi_a(\mathbf{x}) = Z_a/Z$, $Z_a = m_a(\mathbf{x}) \exp(-\lambda_\alpha(\mathbf{x}) p_\alpha(x_a))$, where λ_α ($\alpha = 1,2,\dots,d$) are the Lagrange multipliers associated with the d first-order reproducing conditions, and Z is the [partition function](#). For a smooth prior, the basis functions are also smooth within the convex hull of the nodal set. For a constant prior (state of complete ignorance), H is identical to the Shannon entropy (modulo a constant). From the above expression, the satisfaction of the partition of unity property or the zeroth-order moment constraint ($\sum_a \phi_a = 1$) is evident. On considering the *dual problem* ($\boldsymbol{\lambda}^* = \text{argmin}_{\boldsymbol{\lambda}} \ln Z(\boldsymbol{\lambda})$), well-established numerical algorithms (steepest descent, Newton's method) are utilized to solve the unconstrained convex minimization problem. Once the Lagrange multipliers are determined, the basis functions are computed using the above equation. As with the appeal of radial basis function approximations, here also the spatial dimension does not pose a limitation, since the maximum-entropy formulation and its numerical implementation readily extends to any space-dimension.

Essential Boundary Conditions

Imposition of essential boundary conditions and numerical integration of the Galerkin weak form are the two main chinks in the armor of meshfree methods. In [AO \(2006\)](#), the key properties of convex approximants are established, among which, the facet-reducing-property is pertinent. On any facet (point \mathbf{x} belongs to the facet) of the boundary of the convex hull, only nodes that are located on the facet have non-zero basis function values at \mathbf{x} . This immediately implies that essential boundary conditions can be imposed as in finite elements—note that on weakly convex polygons (polygons with mid-side nodes), interpolation is not met at the middle node. For imposition of essential boundary conditions, cardinality is not a necessary condition. This has not been well-recognized in the meshfree literature, where nodal interpolation through singular weight functions, use of transformations, or other approaches have been pursued. Among the existing techniques to impose essential boundary conditions, Nitsche's method and the blending technique of [Huerta and Fernández-Méndez \(2000\)](#) are promising; use of Lagrange multipliers, modified

variational principles, or techniques that directly couple finite elements to MLS approximations are less appealing within a standard Galerkin method. Imposing linear essential boundary conditions in maximum-entropy meshfree methods can be done as in finite elements for any weight function as a prior (globally- or compactly-supported). This appears to be a simple and elegant means to impose essential boundary conditions in meshfree methods.

Numerical Integration

The issue of essential boundary condition has been discussed, and now the topic of numerical integration is briefly touched upon. If background cells are used within a Galerkin implementation, all the approximation schemes that we have discussed would induce numerical integration errors (with Gauss quadrature) since the intersection of the supports of the basis functions do not coincide with the background cells. Rather than integrating over the precise supports of the basis functions or develop more sophisticated integration rules (both are not very viable alternatives), the development of nodal integration (collocation) schemes is a potentially fruitful direction. Research in stabilized nodal integration techniques for meshfree methods emanated from the work of [Chen et al. \(2001\)](#). In a Lagrangian formulation, on using nodal integration no remapping is required since all quantities are stored at the nodal locations. Large deformation analysis is one of the main application areas where meshfree methods can potentially replace finite elements. The caveat on nodal integration techniques is that ensuring exactness on the patch test alone is insufficient. A better understanding of its relationship with assumed strain methods, stabilization techniques to prevent pressure oscillations, and robust performance in the incompressible limit are needed. Some of these issues are discussed in greater depth for the four-node tetrahedron by [Puso and Solberg \(2005\)](#). Ultimately, for meshfree methods to gain prominence and to reach the mainstream, the conception of nodally integrated stable meshfree (particle) methods is deemed to be critical. All comments and feedback are most welcome.