Representation and approximation of data

Wojciech Czaja and Brian Hunt

February 3, 2015

Outline

Outline

Approximation

The interpretation of polynomials as functions, rather than abstract algebraic objects, forces us to reinterpret the concept and meaning of representations of polynomials as linear combinations of other polynomials. That is, given $f = 3 + x^2 + 17.1x^7$, symbolically we are able to represent it as a linear combination of 3 monomials: 1, x^2, and x^7. In the language of functions, however, equation $f = f(x) = 3 \cdot b_0(x) + b_2(x) + 17.1 \cdot b_7(x)$, with $b_i(x) = x^i$, can be understood very differently, as a set of infinitely many equations for all possible values of x. This is one place where norms, and distances defined by norms, help us with understanding the meaning of such equalities.

Another place, where this geometry can help is when the linear combinations are no longer finite. This is necessarily happening in infinite dimensional vector spaces, and the aforementioned function spaces (such as L^1 or L^2) provide us with situations where there is no simple and intuitive way to think of the infinite sum in the way which we think about finite summations. This is the starting point of the theory of **approximation**.

Interpolation

In addition to approximation, there is also another related concept of **interpolation**. Interpolation is defined as a method of constructing new data points from the set of already known data points. Some well known generic types of methods to achieve this goal are curve fitting or regression analysis.

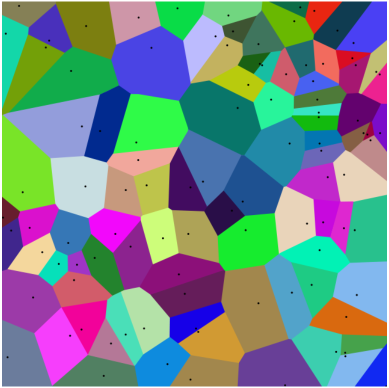
A related (and very important from perspective of modeling) notion is that of **extrapolation**. The difference is rather subtle:

- interpolation is considered when the newly created data points lie within the range of already known values;
- extrapolation pertains to determining the values of points which lie beyond the range of known data.

Norbert Wiener Center for Harmonic Analysis and Applications

Piecewise Constant Interpolation

The simplest method of interpolation is to locate the nearest given data point to the point in question and then to assign the same value. The result is a piecewise constant function. This method is also known as nearest-neighbor interpolation or proximal interpolation.



Norbert Wiener Center for Harmonic Analysis and Applications

Source of imagery: Wikipedia

Interpolation

Other types of interpolation include:

- linear interpolation;
- polynomial interpolation.

Interpolation can be considered a data fit problem. When the function we fit is a polynomial it corresponds to the least squares optimization, with one major difference: the errors on given data **must** be equal to 0. It means that interpolation must be exact on given data, and provides an approximation to the intermediate arguments.
When using polynomials, there is a minimal degree of the polynomial that will depend on the number of data points ($d = n - 1$) and will allow us to fit data exactly.

Norbert Wiener Center for Harmonic Analysis and Applications

Lagrange Interpolation

The question of data interpolation with polynomials is related to the question of representation of polynomials by means other than the monomial basis. Let $f = f(x)$ be any polynomial of degree n, for which we know precisely its values at $n + 1$ distinct points:

$$y_k = f(x_k), \quad \text{for a collection of fixed, distinct points} \quad x_0, \ldots, x_n.$$

Then, we can write f as a linear combination of polynomials λ_k of degree n, $k = 0, \ldots, n$. Indeed, given

$$\lambda_k(x) = \prod_{j \neq k} \frac{x - x_j}{x_k - x_j},$$

we have that

$$f(x) = \sum_{k=0}^{n} y_k \lambda_k(x).$$

(Can you show that λ_k's are indeed linearly independent?)

Chebyshev Polynomials

We define the n^{th} **Chebyshev polynomial** T_n as:

$$T_n(x) = \cos(n \arccos(x)), \quad x \in [-1, 1].$$

(This is indeed a polynomial of degree n with the leading coefficient 2^{n-1}.)
It follows that any polynomial of degree n can be written as a linear combination of Chebyshev polynomials $T_0, \ldots, T_n$.
The intriguing feature of Chebyshev polynomials is that they are orthogonal with respect to the following inner product:

$$\langle p, q \rangle = \sum_{j=0}^{n} p(x_j) q(x_j),$$

for $x_j = \cos\left(\pi \frac{j+0.5}{n+1}\right)$.

Norbert Wiener Center for Harmonic Analysis and Applications

Shannon Sampling Theorem

Interpolation can also be defined in situations where there are infinitely many given data points. It gives us many new possibilities.

- **Cauchy – Kotelnikov – Nyquist – Shannon – Whittaker sampling theorem**: A bandlimited signal can be fully reconstructed if it has been sampled at a rate which exceeds twice the maximum frequency present in the signal.
- Sampled values are assciated with instant times at which they are collected.
- **Borel – Shannon – Whittaker interpolation formula**:

$$f(t) = \sum_{n=-\infty}^{\infty} f(nT)\,\mathrm{sinc}\left(\frac{t-nT}{T}\right).$$

- T is the sampling period and $1/T$ is the sampling rate. The minimum sampling rate which guarantees reconstruction is the **Nyquist rate**.

Data Compression

- Fourier series, bases, and other representation systems generalize the interpolation formula:

$$f(t) = \sum_{n=-\infty}^{\infty} \alpha_n \phi_n(t).$$

- **Sparse representations**: systems which account for most of the information in the signal using only a small number of building blocks. Wavelets, curvelets, contourlets, shearlets, ...
- **Data compression**: retain only the largest coefficients.
- Problems: First sample then compress; Lossy; Computationally expensive.

Norbert Wiener Center for Harmonic Analysis and Applications