Data-dependent and a priori representations

Wojciech Czaja and Brian Hunt

February 12, 2015

Norbert Wiener Center for Harmonic Analysis and Applications

Outline

Norbert Wiener Center for Harmonic Analysis and Applications

Outline

Diversity of basis representations

- Let us start by recalling that a **basis** is a set of linearly independent vectors which can represent every vector in a given vector space through their linear combinations.
- An **orthogonal basis** for a vector space with an inner product, is a basis with vectors which are mutually orthogonal (perpendicular). If the vectors of an orthogonal basis are or length (norm) 1, the resulting basis is an **orthonormal basis (ONB)**.
- Many examples of bases exist in a given Euclidean D-dimensional vector space, ranging from classical $0-1$ bases, through Fourier, Gabor, wavelet, shearlet, curvelet, etc etc.

Norbert Wiener Center for Harmonic Analysis and Applications

Fourier Basis

Given $D > 0$, define the following $D \times D$ matrix:

$$F(m,n) = \frac{1}{\sqrt{N}} e^{2\pi imn/D}, \quad m,n = 0, \dots D-1.$$

The columns (or rows) of this matrix form an orthonormal basis for the space of D-dimensional complex vectors $\mathbb{C}^D$. This basis is called the **Fourier basis**. And the matrix F is known as the Discrete Fourier Transform. Clearly F is a unitary matrix, and as such invertible.

Bases

The role of a basis is to allow us to **represent** elements of the vector space in terms of sequences of scalars (numbers), which are called vector coordinates. This is an important step, because thanks to this **representation**, abstract or complicated objects obtain a uniform mathematical format. The reason for this may not necessarily be clear when we think of the most typical example of a vector space: d-dimensional Euclidean vector space. This is because the Euclidean space is not just a good example of a vector space, it is also a prototypical example, and last but not least - a finite dimensional vector space.
Infinite dimensional vector spaces provide us with more intriguing examples of objects, and the role of a basis which allows us to replace these complicated objects by sequences of numbers becomes much more clear.

- Vector spaces of polynomials;
- Function spaces (Lipschitz, integrable, finite energy functions, etc.)

From data to frame operators

- Given data space X of N vectors $x_n \in \mathbb{R}^D$. Without loss of generality, assume $\sum x_n = 0$ (subtract mean).
- Let P be $D \times N$ matrix whose columns are the data vectors x_n.
- Let $\mathbb{H} = \text{span}\{x_n\}_{n=1}^N \subseteq \mathbb{R}^D$. Define $L : \mathbb{H} \to \mathbb{R}^N$,

$$v \mapsto P^* v = L(v) = \{\langle v, x_n \rangle\} \in \mathbb{R}^N,$$

  and its Hilbert space adjoint $L^* : \mathbb{R}^N \to \mathbb{H} \subseteq \mathbb{R}^D$,

$$w \mapsto L^*(w) = \sum_{n=1}^{N} w[n] x_n, \quad w = (w[1], w[2], \ldots, w[N]).$$

- L is the *Bessel (analysis)* operator, and L^* is the *synthesis* operator.

Norbert Wiener Center for Harmonic Analysis and Applications

The role of covariance

- The frame operator S can be written as

$$S : \mathbb{H} \to \mathbb{H},\ v \mapsto \sum_{n=1}^{N} \langle v, x_n \rangle x_n = (PP^*)v,$$

  where PP^* is $D \times D$.

- Hence, up to a scaling factor and a translation, S is the linear operator identified with the $D \times D$ symmetric covariance matrix $C = \frac{1}{N} PP^*$ of the data space, i.e.

$$C = \frac{1}{N} \left(\sum_{j=1}^{N} x_j[m] x_j[n] \right)_{m,n=1}^{D}, \quad x_j = (x_j[1], \dots, x_j[D]) \in \mathbb{R}^D.$$

Norbert Wiener Center for Harmonic Analysis and Applications

Principal Component Analysis

- The covariance matrix C we have just defined is certainly symmetric and also positive semidefinite, since for every vector y, we have

$$\langle y, Cy \rangle = \frac{1}{N} \sum_{j=1}^{N} |\langle y, x_j \rangle|^2 \geq 0.$$

- Thus, C can be diagonalized, and its eigenvalues are all nonnegative. If K denotes the orthogonal matrix that diagonalizes C, then we have that $K^* CK$ is diagonal and the whole process of analyzing data using the eigenbases of covariance matrix is known as **Principal Component Analysis (PCA)**. K is also known as principal orthogonal decomposition or Karhunen-Loeve transform.
- The columns of K are the eigenvectors of C. The number of positive eigenvalues is the actual number of uncorrelated parameters, or degrees of freedom in the original data set X. Each eigenvalue is the variance of its degree of freedom.