MATHEMATICAL DESCRIPTIONS OF IMAGING SYSTEMS

Harrison H. Barrett
University of Arizona

OBJECTIVES:

Introduce a vector-space formalism for describing imaging systems

Review some familiar kinds of linear systems (LSIV, LSV, matrices)

Introduce the not-so-familiar math that describes real-world digital imaging systems

Discuss special problems such as null functions that are inevitable with digital systems

OUTLINE

- Objects and images as vectors
- Scalar products and Hilbert spaces
- Imaging systems as mappings between vector spaces
- Categories of linear imaging systems
- Some familiar mathematical descriptions of linear systems
- The real world: Continuous-to-discrete mappings
- Null functions and measurement functions of digital imaging systems

REFERENCE

H. H. Barrett and K. J. Myers, Foundations of Image Science (Wiley, 2004)

Chap. 1: Vectors and Operators

Chap. 7: Deterministic Descriptions of Imaging Systems

OBJECTS

An object to be imaged is a spatial or spatio-temporal distribution of some physical quantity.

Examples:

- Spatial distribution of reflectance on a 2D plane

- 3D distribution of fluorophors in fluorescence microscopy

- Radiance of the night sky (scalar-valued function of 4 variables)

- Electric field emerging from a coherently illuminated transparency (vector-valued function of two spatial variables)

- Strain wave produced by an earthquake in seismic imaging (tensor-valued function of 3 spatial and one temporal coordinate)

NOTATION FOR OBJECTS

Key point: Real-world objects are always *functions*. For simplicity, we shall consider only scalar-valued functions.

We shall use the symbol $f(\,\cdot\,)$ to represent an object, and the argument of the function will be one or more spatial coordinates and possibly the time.

Thus the reflectance of a static 2D surface would be written as $f(x, y)$, and the time-varying distribution of fluorophores in microscopy would be $f(x, y, z, t)$.

A useful shorthand is to use a vector $\mathbf{r}$ as the argument of $f(\,\cdot\,)$ without being too specific about the meaning of $\mathbf{r}$. Thus $f(\mathbf{r})$ will mean $f(x, y)$ for a 2D static object and $f(x, y, z, t)$ for a 3D dynamic object. When necessary, we shall use the symbol q to specify the dimensionality of $\mathbf{r}$. Thus $q = 2$ if $\mathbf{r} = (x, y)$ and $q = 4$ if $\mathbf{r} = (x, y, z, t)$.

OBJECTS AS VECTORS

A simple but general definition of a *vector* is that it is something that can be multipled by scalars and added to other vectors of the same ilk.

In ordinary 3D vector analysis, $\mathbf{A} \equiv (A_1, A_2, A_3)$ is a vector since:

$$\alpha\mathbf{A} = (\alpha A_x, \alpha A_y, \alpha A_z)\,, \qquad \mathbf{A} + \mathbf{B} = (A_1 + B_1, A_2 + B_2, A_3 + B_3)\,,$$

where α is an arbitrary scalar.

Functions can be regarded as vectors also since $\alpha f(\mathbf{r})$ and $f_i(\mathbf{r}) + f_j(\mathbf{r})$ are well-defined (where $f_i(\mathbf{r})$ and $f_j(\mathbf{r})$ are two different functions of the same variables).

The collection of all possible vectors of a given kind is called a *vector space*. For example, the vector space of ordinary 3D vector analysis is the set of all possible real numbers A_1, A_2 and A_3. Since three numbers suffice to specify any vector $\mathbf{A}$, this vector space is three-dimensional; function spaces are often infinite-dimensional.

Norms and scalar products

In ordinary 3D vector analysis, the length or *norm* of a vector is defined by

$$||\mathbf{A}|| = \sqrt{A_1^2 + A_2^2 + A_3^2}\,,$$

and a scalar or dot product between two vectors is defined by

$$(\mathbf{A}, \mathbf{B}) \equiv \mathbf{A} \cdot \mathbf{B} = A_1 B_1 + A_2 B_2 + A_3 B_3\,.$$

Note that

$$||\mathbf{A}||^2 = (\mathbf{A}, \mathbf{A})\,.$$

If we generalize to N-dimensional vectors and allow the components to be complex, we can define

$$(\mathbf{A}, \mathbf{B}) \equiv \sum_{n=1}^{N} A_n^* B_n\,,$$

$$||\mathbf{A}||^2 = \sum_{n=1}^{N} A_n^* A_n = \sum_{n=1}^{N} |A_n|^2\,.$$

NORMS AND SCALAR PRODUCTS OF FUNCTIONS

Repeating from the last page:

$$(\mathbf{A}, \mathbf{B}) \equiv \sum_{n=1}^{N} A_n^* B_n \, ,$$

$$||\mathbf{A}||^2 = \sum_{n=1}^{N} A_n^* A_n = \sum_{n=1}^{N} |A_n|^2 \, .$$

Now we can generalize to functions just by replacing the sums by integrals. For example, if we denote the 1D functions $f_i(x)$ and $f_j(x)$ as $\mathbf{f}_i$ and $\mathbf{f}_j$, respectively, when we want to think of them as vectors, we can define

$$(\mathbf{f}_i, \, \mathbf{f}_j) \equiv \int_{-\infty}^{\infty} dx \, f_i^*(x) \, f_j(x) \, ,$$

$$||\mathbf{f}_i||^2 \equiv \int_{-\infty}^{\infty} dx \; |f_i(x)|^2 \, .$$

Compare:

$$(\mathbf{A}, \mathbf{B}) \equiv \sum_{n=1}^{N} A_n^* B_n \,, \qquad (\mathbf{f}_i, \mathbf{f}_j) \equiv \int_{-\infty}^{\infty} dx\, f_i^*(x)\, f_j(x)\,,$$

In a very loose sense, the continuous variable x has replaced the discrete index n, and the space of all functions $f(x)$ for which $||\mathbf{f}||^2 < \infty$ is an infinite-dimensional vector space (made up of one-dimensional functions!).

A vector space in which a norm is defined is called – what else? – a *normed vector space* (or with a few other technicalities, a *Banach space*), and a Banach space in which a scalar product is defined is called a *Hilbert space*.

The particular Hilbert space defined by the norm and scalar product just given is called $\mathbb{L}_2(\mathbb{R})$. Here, the $\mathbb{L}_2$ part tells you that it is a space of square-integrable functions, and $\mathbb{R}$ tells you that the integral is over the real line, $-\infty < x < \infty$.

OBJECT SPACE

Objects are functions $f(\mathbf{r})$ where $\mathbf{r}$ is a qD vector, and there is surely no difficulty in assuming that all real-world objects have finite "energy", so

$$||\mathbf{f}\,||^2 \equiv \int_\infty d^q r\,|f(\mathbf{r})|^2 < \infty\,,$$

where the funny ∞ on the integral sign just means to integrate over the infinite range of all q variables. For example, if $q = 2$, the condition above is

$$||\mathbf{f}\,||^2 \equiv \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\,|f(x,y)|^2 < \infty\,.$$

If we define the scalar product for arbitrary q by

$$(\mathbf{f}_i,\,\mathbf{f}_j) \equiv \int_\infty d^q r f_i^*(\mathbf{r})\, f_j(\mathbf{r})\,,$$

then the vector space is denoted $\mathbb{L}_2(\mathbb{R}^q)$. Other variables and other ranges of integration are possible, but we shall always assume that an object is a square-integrable function over some domain (hence a vector in some Hilbert space), and we shall denote that space in general as $\mathbb{U}$.

DISCRETE REPRESENTATIONS OF OBJECTS

Though real objects are functions in an infinite-dimensional Hilbert space, we often approximate them as finite sums of *expansion functions* such as pixels:

$$f_a(\mathbf{r}) = \sum_{n=1}^{N} \theta_n \, \phi_n(\mathbf{r}) \, ,$$

where subscript a denote *approximate*

- The subscript is crucial; we must not confuse $f_a(\mathbf{r})$ and $f(\mathbf{r})$.

- If the expansion functions $\phi_n(\mathbf{r})$ are square-integrable, then $f_a(\mathbf{r})$ lies in a finite-dimensional subspace of $\mathbb{U}$; we call this subspace *representation space*.

- The coefficients in the expansion, the set $\{\theta_n, \, n = 1, ..., N\}$, can be regarded as an N-dimensional vector denoted $\boldsymbol{\theta}$; if the expansion functions are known, this vector is an equivalent way of specifying the approximate representation $f_a(\mathbf{r})$.

IMAGES

Images are either analog or digital.

An analog image, such as the irradiance in the focal plane of an optical system, is a function, denoted $g(\mathbf{r}_d)$, where the subscript d, indicating the *detector* plane, distinguishes $\mathbf{r}_d$ from the vector $\mathbf{r}$ in the functional description of the object.

The dimensionality of $\mathbf{r}_d$ is denoted s, and it need not be the same as the dimensionality q of $\mathbf{r}$; for example, our eyes map the 3D world to our 2D retina, so $q = 3$ and $s = 2$

A digital image, such as the output from a CCD camera, is not a function. Instead it is a discrete set of M numbers $\{g_m, m = 1, ...M\}$. The whole set can be denoted as $\mathbf{g}$.

COMMENT

The term digital image may also suggest that the grey values are discrete, say quantized on an 8-bit scale (0 - 255), or the gray values may be integers when we consider a photon-counting detector. In the math that follows, we do not need to make this distinction: g represents a discrete set of numbers, but the *values* of these numbers may be continuous or discrete.

IMAGES AS VECTORS

Digital images are easily recognized as MD vectors. A scalar product between two different digital images (with the same M) is given by

$$(\mathbf{g}_i, \mathbf{g}_j) \equiv \sum_{m=1}^{M} g_{im}^* \, g_{jm} \, .$$

This scalar product defines the MD *Euclidean space* denoted $\mathbb{E}^M$.

Analog images can be regarded as vectors in an infinite-dimensional Hilbert space, with scalar product defined by

$$(\mathbf{g}_i, \, \mathbf{g}_j) \equiv \int_\infty d^s r_d \, g_i^*(\mathbf{r}_d) \, g_j(\mathbf{r}_d) \, .$$

This vector space is denoted $\mathbb{L}_2(\mathbb{R}^s)$.

In general, we denote image space as $\mathbb{V}$, recognizing that it is finite-dimensional for digital images and infinite-dimensional for analog ones.

IMAGING AS MAPPING

So far we know that:

Objects are vectors in the infinite-dimensional Hilbert space $\mathbb{U}$

Images are vectors in the Hilbert space $\mathbb{V}$, which is finite-dimensional for digital images and infinite-dimensional for analog ones.

Thus an imaging system *maps* or *transduces* an object to an image. In the absence of noise, the image is related to the object by

$$\mathbf{g} = \boldsymbol{\mathcal{H}}\mathbf{f} \,,$$

where $\boldsymbol{\mathcal{H}}$ is an operator describing the imaging system. The mathematicians would write

$$\boldsymbol{\mathcal{H}} : \mathbb{U} \rightarrow \mathbb{V} \,,$$

meaning that $\boldsymbol{\mathcal{H}}$ operates on a vector in $\mathbb{U}$ (i. e., an object) and produces a vector in $\mathbb{V}$ (an image). $\mathbb{U}$ is the *domain* of $\boldsymbol{\mathcal{H}}$, and $\mathbb{V}$ is its *range*.

LINEARITY

Today we will consider only *linear* imaging systems. A linear operator is one that satisfies

$$\boldsymbol{\mathcal{H}}(\alpha \mathbf{f}_1 + \beta \mathbf{f}_2) = \alpha \mathbf{g}_1 + \beta \mathbf{g}_2 \,,$$

where α and β are arbitrary scalars and

$$\boldsymbol{\mathcal{H}}\mathbf{f}_1 = \mathbf{g}_1 \text{ and } \boldsymbol{\mathcal{H}}\mathbf{f}_2 = \mathbf{g}_2 \,.$$

CATEGORIES OF OPERATORS

We can distinguish four general categories of operators, depending on whether the domain $\mathbb{U}$ and the range $\mathbb{V}$ are finite-dimensional or infinite dimensional.

- Continuous-to-continuous (CC) operators map functions of continuous variables to other functions of continuous variables.

- Discrete-to-discrete (DD) operators map discrete vectors (finite sets of numbers) to other discrete vectors.

- Continuous-to-discrete (CD) operators map functions of continuous variables to discrete vectors.

- Discrete-to-continuous (DC) operators map functions of continuous variables to discrete vectors.

EXAMPLES FROM IMAGING

- Continuous-to-continuous (CC) operators

  Lenses, mirrors, etc.
  Computed tomography system, including reconstruction and display

- Discrete-to-discrete (DD) operators

  Matrix model for an imaging system

- Continuous-to-discrete (CD) operators

  CCD camera
  Computed tomography, mapping from object to raw projection data

- Discrete-to-continuous (DC) operators

  Display
  Computer-generated holography

Linear continuous-to-continuous mappings

The most general form of a linear CC mapping from $\mathbb{L}_2(\mathbb{R}^q)$ to $\mathbb{L}_2(\mathbb{R}^s)$ is

$$g(\mathbf{r}_d) = \int_\infty d^q r \, h(\mathbf{r}_d, \mathbf{r}) \, f(\mathbf{r}) \, .$$

The kernel $h(\mathbf{r}_d, \mathbf{r})$ is a very general kind of *point response function* (PRF). It specifies how a point at some location, say $\mathbf{r} = \mathbf{r}_0$, produces an image at a point $\mathbf{r}_d$ in the image (detector) plane.

If the input to the system is a qD Dirac delta function, $f(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r_0})$, then the output is

$$g(\mathbf{r}_d) = \int_\infty d^q r \, h(\mathbf{r}_d, \mathbf{r}) \, \delta(\mathbf{r} - \mathbf{r}_0) = h(\mathbf{r}_d, \mathbf{r}_0) \, .$$

Note that $h(\mathbf{r}_d, \mathbf{r}_0)$ is a function of $q + s$ variables. Even a simple lens mapping one plane to another has a PRF that depends on four variables (x, y, x_d, y_d).

Linear shift-invariant systems

Sometimes we consider $q = s$ and assume that a CC system is *shift-invariant*, in which case we can write the mapping as a *convolution*:

$$g(\mathbf{r}_d) = \int_\infty d^q r \, h(\mathbf{r}_d - \mathbf{r}) \, f(\mathbf{r}) \, .$$

Again the kernel is the PRF, but now it depends on only q variables (e.g. 2 variables instead of 4 when imaging from one plane to another). If the input to a linear shift-invariant (LSIV) system is a qD delta function, $f(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r_0})$, then the output is

$$g(\mathbf{r}_d) = \int_\infty d^q r \, h(\mathbf{r}_d - \mathbf{r}) \, \delta(\mathbf{r} - \mathbf{r}_0) = h(\mathbf{r}_d - \mathbf{r}_0) \, .$$

Thus the qD function $h(\mathbf{r}_d)$ suffices to specify the PRF for all $\mathbf{r}$; it just gets shifted in the image plane as the object point shifts.

Caution: A linear shift-invariant system is a mythical beast.

Linear discrete-to-discrete mappings

For DD operators, input and output are vectors. An ND vector is a set of N numbers which can be arranged as an $N \times 1$ column vector, or equivalently it can be regarded as a vector in the Euclidean space $\mathbb{E}^N$. An example in imaging is the set of coefficients $\{\theta_n, n = 1, ..., N\}$ that specifies the approximate object representation $f_a(\mathbf{r})$.

The most general linear mapping from $\mathbb{E}^N$ to $\mathbb{E}^M$ is the $M \times N$ matrix (M rows and N columns) denoted $\mathbf{H}$:

$$\mathbf{g} = \mathbf{H}\boldsymbol{\theta} \qquad \text{or} \qquad g_m = \sum_{n=1}^{N} H_{mn}\theta_n \,,$$

where

$$\mathbf{g} : M \times 1 \,, \qquad \mathbf{H} : M \times N \,, \qquad \boldsymbol{\theta} : N \times 1 \,.$$

PRF for linear DD systems

To repeat from the last slide,

$$g_m = \sum_{n=1}^{N} H_{mn}\theta_n \,.$$

Let $\theta_n = \delta_{nn_0}$ (Kronecker delta, not Dirac), so that $\boldsymbol{\theta}$ is an $N \times 1$ column vector with a 1 in row n_0 and 0 everywhere else. Then

$$g_m = \sum_{n=1}^{N} H_{mn}\, \delta_{nn_0} = H_{mn_0} \,.$$

Thus a column of $\mathbf{H}$ is the PRF for a DD system.

Linear continuous-to-discrete mappings

For a CD imaging system, objects are functions $f(\mathbf{r})$, where $\mathbf{r}$ is a q-dimensional vector (usually $q = 2$ or 3). The images are discrete vectors, $\mathbf{g} = \{g_m, m = 1, ..., M\}$.

Linear CD systems have the general form

$$g_m = \int_{\mathbf{S}_f} d^q r \, h_m(\mathbf{r}) \, f(\mathbf{r})$$

or in operator form

$$\mathbf{g} = \boldsymbol{\mathcal{H}}\mathbf{f}$$

where $\boldsymbol{\mathcal{H}}$ is a linear continuous-to-discrete (CD) operator.

PRF for linear CD systems

Repeating:

$$g_m = \int_{\mathbf{S}_f} d^q r \, h_m(\mathbf{r}) \, f(\mathbf{r})$$

Once again the kernel $h_m(\mathbf{r})$ is the PRF of the system. For a delta-function input, the output is

$$g_m = \int_{\mathbf{S}_f} d^q r \, h_m(\mathbf{r}) \, \delta(\mathbf{r} - \mathbf{r}_0) = h_m(\mathbf{r}_0) \, .$$

Another name for $h_m(\mathbf{r})$ is the *sensitivity function* since it specifies how sensitive the m^{th} discrete detector element is to a point source at $\mathbf{r} = \mathbf{r}_0$.

KEY POINT: All CD operators have null functions:

$$\mathcal{H}\mathbf{f}_{null} = 0$$

You can never recover the object $f(\mathbf{r})$ exactly, even in the absence of noise.

An arbitrary object can be decomposed uniquely into its null and measurement components:

$$\mathbf{f} = \mathbf{f}_{meas} + \mathbf{f}_{null} \qquad \text{or} \qquad f(\mathbf{r}) = f_{meas}(\mathbf{r}) + f_{null}(\mathbf{r})\,,$$

and only the measurement component $f_{meas}(\mathbf{r})$ contributes to the measured data.

STILL TO COME

In coming lectures we will learn how to:

- Characterize the null space of arbitrary linear operators
- Decompose any object into measurement and null components
- Develop inversion methods for recovering the measurement part in the absence of noise
- Learn how to deal with noisy data

Our main tools will be singular-value decomposition and the Moore-Penrose pseudoinverse.