MMAT5390: Mathematical Image Processing

Chapter 1: Basic concepts in Digital Image Processing

Mathematical imaging aims to develop mathematical models to process a digital image. The main tasks include enhancing the visual quality of a corrupted image and extracting important information from an image for the purpose of image understanding. Most mathematical models are done by transforming one image into another or by decomposing an image into meaningful components. In this chapter, we will explain some basic concepts in mathematical image processing. The definition of a digital image will firstly be described. The basic idea of image transformation and image decomposition will then be described in details. Finally, various measures to quantify the similarity between images will be explained.

1 Definition of digital images

A digital image captures the brightness at each pixel and represented by a numerical value, called the *pixel value*. Mathematically, a digital image can be understood as a matrix, which is array of numbers recording the pixel values at each pixels.

Definition 1.1. A digital **image** of width m pixels and height n pixels can be represented by a matrix $I \in \mathbb{R}^{n\times m}$. Mathematically, I belongs to the subset $\mathcal{I}$ of $n \times m$ matrices:

$$\mathcal{I} = \{I \in \mathbb{R}^{n\times m} \ : \ 0 \leq I(i,j) \leq R \ \text{ for } \ 1 \leq i \leq n, 1 \leq j \leq m\}$$

Typical values of the upper bound R of pixel values include 1 for greyscale images, and 255 for Red-Green-Blue (RGB) or Cyan-Magenta-Yellow-blacK (CMYK) images.

The main idea of mathematical imaging can be described as follows:

1. Given a noisy/distorted image $f \in \mathcal{I}$, find a suitable image transformation $T : \mathcal{I} \to \mathcal{I}$ such that $g := T(f)$ is the restored (good) image.
2. Given a distorted image $g \in \mathcal{I}$. We assume g is distorted by an image transformation $T : \mathcal{I} \to \mathcal{I}$ and corrupted by some noise n. Mathematically, we can write:

   $$g = T(f) + n,$$

   where f is the unknown good image. Given g and T, our goal is to find the good image f and n. This kind of problems is called the *inverse problem*. Mathematical imaging is often considered as an inverse problem.

2 Basic idea of image transformation

We will first focus on the linear image transformation $\mathcal{O}$. For simplicity, we will assume images in $\mathcal{I}$ consist of square images of size N. In other words, we assume $m = n = N$.

Definition 2.1. A image transformation $\mathcal{O} : \mathcal{I} \to \mathcal{I}$ is **linear** if it satisfies:

$$\mathcal{O}(af + bg) = a\mathcal{O}(f) + b\mathcal{O}(g)$$

for all $f, g \in \mathcal{I}$ and $a, b \in \mathbb{R}$.

Take $f \in \mathcal{I}$. Let

$$f = \begin{pmatrix} f(1,1) & f(1,2) & \cdots & f(1,N) \\ f(2,1) & f(2,2) & \cdots & f(2,N) \\ \vdots & \vdots & \ddots & \vdots \\ f(N,1) & f(N,2) & \cdots & f(N,N) \end{pmatrix} = \sum_{i=1}^{N}\sum_{j=1}^{N} \begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & f(i,j) & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}$$

with $f(i,j)$ at i-th row and j-th column.
Let $g = \mathcal{O}(f)$. Since $\mathcal{O}$ is linear,

$$g(\alpha,\beta) = \sum_{x=1}^{N}\sum_{y=1}^{N} f(x,y)h(x,\alpha,y,\beta)$$

where

$$h(x,\alpha,y,\beta) = \left[\mathcal{O}\left(\begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}\right)\right]_{\alpha,\beta}$$

with 1 at x-th row and y-th column.

Remark. *$h(x,\alpha,y,\beta)$ determines how much the input value at (x,y) influences the output value at (α,β).*

Definition 2.2. $h(x,\alpha,y,\beta)$ is usually called the **point spread function (PSF)**.

In mathematical imaging, two types of linear image transformations are particularly useful. They are, namely, the *separable* and *shift-invariant* linear image transformations.

Definition 2.3. The PSF $h(x,\alpha,y,\beta)$ is called **shift-invariant** if there exists a function $\tilde{h}$ such that

$$h(x,\alpha,y,\beta) = \tilde{h}(\alpha - x, \beta - y)$$

for all $1 \le x, y, \alpha, \beta \le N$.

Definition 2.4. The PSF is called **separable** if there exist functions h_c and h_r such that

$$h(x,\alpha,y,\beta) \equiv h_c(x,\alpha)h_r(y,\beta)$$

for all $1 \le x, y, \alpha, \beta \le N$.

Some interesting properties related to shift-invariant and separable linear image transformation can be observed. To begin with, a shift-invariant linear image transformation is related to something called the *convolution*.

Definition 2.5. Consider two digital images $f \in \mathcal{I}$ and $g \in \mathcal{I}$. Assume that they are periodically extended, that is,

$$f(x,y) = f(x+pN, y+qN), g(x,y) = g(x+pN, y+qN)$$

where p and q are any integers. The **convolution** $f * g$ of two images $f \in \mathcal{I}$ and $g \in \mathcal{I}$ is defined as

$$f * g(\alpha,\beta) = \sum_{x=1}^{N}\sum_{y=1}^{N} f(x,y)g(\alpha - x, \beta - y).$$

Obviously, if the PSF h of a linear image transformation $\mathcal{O}$ is shift-invariant, then the image transformation is a convolution because

$$\mathcal{O}(f)(\alpha,\beta) = \sum_{x=1}^{N}\sum_{y=1}^{N} f(x,y)h(\alpha - x, \beta - y).$$

Remark 1.6.

- If PSF h is separable, then

$$\mathcal{O}(f)(\alpha,\beta) = \sum_{x=1}^{N} h_c(x,\alpha) \sum_{y=1}^{N} f(x,y) h_r(y,\beta),$$

  which consists of two one-dimensional linear transformations.

- If PSF h is both separable and shift-invariant, then

$$\mathcal{O}(f)(\alpha,\beta) = \sum_{x=1}^{N} h_c(\alpha - x) \sum_{y=1}^{N} f(x,y) h_r(\beta - y),$$

  which consist of two one-dimensional convolutions.

Conventionally, the linear image transformation can be represented by a big matrix H. In other words, the linear image transformation can be considered as a multiplication of the image by a big matrix H.
Note that:

$$\begin{aligned} g(\alpha,\beta) = {} & f(1,1)h(1,\alpha,1,\beta) + f(2,1)h(2,\alpha,1,\beta) + \cdots + f(N,1)h(N,\alpha,1,\beta) \\ & + f(1,2)h(1,\alpha,2,\beta) + \cdots + f(N,2)h(N,\alpha,2,\beta) + \cdots \\ & + f(1,N)h(1,\alpha,N,\beta) + \cdots + f(N,N)h(N,\alpha,N,\beta) \end{aligned}$$

Rewrite $g(\alpha,\beta) = \vec{h}^T_{\alpha\beta} \cdot \vec{f}^T$ where

$$\begin{aligned} \vec{h}^T_{\alpha\beta} & \equiv [h(1,\alpha,1,\beta), \cdots, h(N,\alpha,1,\beta), h(1,\alpha,2,\beta), \cdots, h(N,\alpha,2,\beta), \\ & \quad \cdots, h(N,\alpha,N,\beta)] \\ \vec{f}^T & \equiv [f(1,1), \cdots, f(N,1), f(1,2), \cdots, f(N,2), \cdots, f(N,N)] \end{aligned}$$

Let $\vec{g}^T \equiv [g(1,1), \cdots, g(N,1), g(1,2), \cdots, g(N,2), \cdots, g(N,N)]$, then $\vec{g} = H\vec{f}$ where

$$H = \begin{pmatrix} \left(\begin{array}{cc} & x \to \\ \alpha\downarrow & \begin{pmatrix} y=1 \\ \beta=1 \end{pmatrix} \end{array}\right) & \left(\begin{array}{cc} & x \to \\ \alpha\downarrow & \begin{pmatrix} y=2 \\ \beta=1 \end{pmatrix} \end{array}\right) & \cdots & \left(\begin{array}{cc} & x \to \\ \alpha\downarrow & \begin{pmatrix} y=N \\ \beta=1 \end{pmatrix} \end{array}\right) \\ \left(\begin{array}{cc} & x \to \\ \alpha\downarrow & \begin{pmatrix} y=1 \\ \beta=2 \end{pmatrix} \end{array}\right) & \left(\begin{array}{cc} & x \to \\ \alpha\downarrow & \begin{pmatrix} y=2 \\ \beta=2 \end{pmatrix} \end{array}\right) & \cdots & \left(\begin{array}{cc} & x \to \\ \alpha\downarrow & \begin{pmatrix} y=N \\ \beta=2 \end{pmatrix} \end{array}\right) \\ \vdots & \vdots & & \vdots \\ \left(\begin{array}{cc} & x \to \\ \alpha\downarrow & \begin{pmatrix} y=1 \\ \beta=N \end{pmatrix} \end{array}\right) & \left(\begin{array}{cc} & x \to \\ \alpha\downarrow & \begin{pmatrix} y=2 \\ \beta=N \end{pmatrix} \end{array}\right) & \cdots & \left(\begin{array}{cc} & x \to \\ \alpha\downarrow & \begin{pmatrix} y=N \\ \beta=N \end{pmatrix} \end{array}\right) \end{pmatrix}$$

Definition 2.6. H is called the **transformation matrix** of $\mathcal{O}$.

Example 2.7. A linear operator is such that it replaces the value of each pixel by the average of its four nearest neighbours. Assume the image is repeated in all directions. Apply this operator $\mathcal{O}$ to a 3×3 image. Find the transformation matrix corresponding to $\mathcal{O}$.

Solution. The 3×3 image looks like

$$\begin{array}{c|ccc|c} f_{33} & f_{31} & f_{32} & f_{33} & f_{31} \\ \hline f_{13} & f_{11} & f_{12} & f_{13} & f_{11} \\ f_{23} & f_{21} & f_{22} & f_{23} & f_{21} \\ f_{33} & f_{31} & f_{32} & f_{33} & f_{31} \\ \hline f_{13} & f_{11} & f_{12} & f_{13} & f_{11} \end{array}$$

$$\therefore g_{11} = \frac{f_{21} + f_{31} + f_{12} + f_{13}}{4}$$

$$g_{21} = \frac{f_{11} + f_{31} + f_{22} + f_{23}}{4},$$

and so on.

By a simple checking, we observe that the transformation matrix H can be written as

$$H = \begin{bmatrix} 0 & 1/4 & 1/4 & 1/4 & 0 & 0 & 1/4 & 0 & 0 \\ 1/4 & 0 & 1/4 & 0 & 1/4 & 0 & 0 & 1/4 & 0 \\ 1/4 & 1/4 & 0 & 0 & 0 & 1/4 & 0 & 0 & 1/4 \\ 1/4 & 0 & 0 & 0 & 1/4 & 1/4 & 1/4 & 0 & 0 \\ 0 & 1/4 & 0 & 1/4 & 0 & 1/4 & 0 & 1/4 & 0 \\ 0 & 0 & 1/4 & 1/4 & 1/4 & 0 & 0 & 0 & 1/4 \\ 1/4 & 0 & 0 & 1/4 & 0 & 0 & 0 & 1/4 & 1/4 \\ 0 & 1/4 & 0 & 0 & 1/4 & 0 & 1/4 & 0 & 1/4 \\ 0 & 0 & 1/4 & 0 & 0 & 1/4 & 1/4 & 1/4 & 0 \end{bmatrix}.$$

Example 2.8. Consider an image transformation on a 2×2 image. Suppose the matrix representation of the image transformation is given by:

$$H = \begin{pmatrix} 2 & 0 & 1 & 0 \\ 4 & 2 & 2 & 1 \\ 3 & 0 & 4 & 0 \\ 6 & 3 & 8 & 4 \end{pmatrix}.$$

Prove that the image transformation is separable. Find g_1 and g_2 such that:

$$h(x, \alpha, y, \beta) = g_1(x, \alpha) g_2(y, \beta).$$

Solution. Recall that:

$$H = \begin{pmatrix} \alpha \downarrow \overset{x \rightarrow}{\begin{pmatrix} y = 1 \\ \beta = 1 \end{pmatrix}} & \alpha \downarrow \overset{x \rightarrow}{\begin{pmatrix} y = 2 \\ \beta = 1 \end{pmatrix}} \\ \alpha \downarrow \overset{x \rightarrow}{\begin{pmatrix} y = 1 \\ \beta = 2 \end{pmatrix}} & \alpha \downarrow \overset{x \rightarrow}{\begin{pmatrix} y = 2 \\ \beta = 2 \end{pmatrix}} \end{pmatrix}$$

If H is separable, then $h(x, \alpha, y, \beta) = g_1(x, \alpha) g_2(y, \beta)$ for some g_1 and g_2. Then,

$$H = \begin{pmatrix} h(1,1,1,1) & h(2,1,1,1) & h(1,1,2,1) & h(2,1,2,1) \\ h(1,2,1,1) & h(2,2,1,1) & h(1,2,2,1) & h(2,2,2,1) \\ h(1,1,1,2) & h(2,1,1,2) & h(1,1,2,2) & h(2,1,2,2) \\ h(1,2,1,2) & h(2,2,1,2) & h(1,2,2,1) & h(2,2,2,2) \end{pmatrix}.$$

We can easily check that

$$H = \begin{pmatrix} g_2(1,1)G_1 & g_2(2,1)G_1 \\ g_2(1,2)G_1 & g_2(2,2)G_1 \end{pmatrix},$$

where

$$G_1 = \begin{pmatrix} g_1(1,1) & g_1(2,1) \\ g_1(1,2) & g_1(2,2) \end{pmatrix}.$$

In our case,

$$H = \begin{pmatrix} 2G_1 & 1G_1 \\ 3G_1 & 4G_1 \end{pmatrix},$$

where

$$G = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}.$$

Thus, $g_1(1,1) = 1$, $g_1(2,1) = 0$, $g_1(1,2) = 2$ and $g_1(2,2) = 1$. Similarly, $g_2(1,1) = 2$, $g_2(2,1) = 1$, $g_2(1,2) = 3$ and $g_2(2,2) = 4$.

Example 2.9. Suppose $H \in \mathbb{R}^{4\times 4}$ is applied to a 2×2 image. Let

$$H = \begin{pmatrix} 1 & 1 & 1 & 2 \\ 2 & 3 & 2 & 4 \\ 2 & 1 & 1 & 1 \\ 6 & 1 & 1 & 1 \end{pmatrix}.$$

Is H shift-invariant?

Solution. We can easily check that $h(1,2,1,1) = 2$ and $h(1,2,2,2) = 1$. Hence, H is not shift-invariant.

3 Properties of shift-invariant/separable image transformation

Properties of shift-invariant image transformation

Definition 3.1. The **circulant matrix** $V := circ(\mathbf{v})$ associated to the vector $\mathbf{v} = (v_0, v_1, \ldots, v_{n-1})^T \in \mathbb{C}^n$ is an $n \times n$ matrix whose columns are given by iterations of shift-operator T acting on $\mathbf{v}$, where $T : \mathbb{C}^n \to \mathbb{C}^n$ is given by:

$$T(\begin{pmatrix} v_0 \\ v_1 \\ \vdots \\ v_{n-1} \end{pmatrix}) = \begin{pmatrix} v_{n-1} \\ v_0 \\ \vdots \\ v_{n-2} \end{pmatrix}$$

Hence, the k-th column of V is given by $T^{k-1}(\mathbf{v})$ $(k = 1, 2, \ldots, n)$. In other words,

$$V = \begin{pmatrix} v_0 & v_{n-1} & \cdots & v_1 \\ v_1 & v_0 & \cdots & v_2 \\ \vdots & \vdots & \cdots & \vdots \\ v_{n-1} & v_{n-2} & \cdots & v_0 \end{pmatrix}.$$

Definition 3.2. A matrix V is called **block-circulant** if it is of the following form:

$$V = \begin{pmatrix} H_0 & H_{n-1} & \cdots & H_1 \\ H_1 & H_0 & \cdots & H_2 \\ \vdots & \vdots & \cdots & \vdots \\ H_{n-1} & H_{n-2} & \cdots & H_0 \end{pmatrix}$$

where each H_i is a circulant matrix $(i = 0, 1, 2, \ldots, n-1)$.

Theorem 3.3. *Let H be the transformation matrix of a shift-invariant linear image transformation on $M_{N\times N}(\mathbb{R})$, with h_s being N-periodic in both arguments. Suppose*

$$H = \begin{pmatrix} A_{11} & \cdots & A_{1N} \\ \vdots & \cdots & \vdots \\ A_{N1} & \cdots & A_{NN} \end{pmatrix}$$

where each A_{ij} is an $N \times N$ block matrix. Then, each A_{ij} is a circulant matrix.

Proof. Note that

$$A_{ij} = \begin{pmatrix} & x \to \\ \alpha \downarrow & \begin{pmatrix} y = j \\ \beta = i \end{pmatrix} \end{pmatrix}.$$

Thus,

$$A_{ij} = \begin{pmatrix} h(1,1,j,i) & h(2,1,j,i) & \cdots & h(N,1,j,i) \\ h(1,2,j,i) & h(2,2,j,i) & \cdots & h(N,2,j,i) \\ \vdots & & \vdots & \\ h(1,N,j,i) & h(2,N,j,i) & \cdots & h(N,N,j,i) \end{pmatrix}.$$

By assumption, we have $h(x, \alpha, y, \beta) = h_s(\alpha - x, \beta - y)$ with h_s being N-periodic in both arguments. We conclude that

$$A_{ij} = \begin{pmatrix} h_s(0, i-j) & h_s(-1, i-j) & \cdots & h_s(1-N, i-j) \\ h_s(1, i-j) & h_s(0, i-j) & \cdots & h_s(2-N, i-j) \\ \vdots & & \vdots & \\ h_s(N-1, i-j) & h_s(N-2, i-j) & \cdots & h_s(0, i-j) \end{pmatrix}$$
$$= \begin{pmatrix} h_s(0, i-j) & h_s(N-1, i-j) & \cdots & h_s(1, i-j) \\ h_s(1, i-j) & h_s(0, i-j) & \cdots & h_s(2, i-j) \\ \vdots & & \vdots & \\ h_s(N-1, i-j) & h_s(N-2, i-j) & \cdots & h_s(0, i-j) \end{pmatrix}$$

which is circulant. □

Theorem 3.4. *Under the same setup as in* **Theorem 3.3**, *H is block-circulant.*

Proof. Exercise.

For more details, see Appendix.

Properties of separable image transformation

Recall: Separable h means $h(x, \alpha, y, \beta) = h_c(x, \alpha) h_r(y, \beta)$. Then, if $g = Hf$, we have

$$g(\alpha, \beta) = \sum_{x=1}^{N} h_c(x, \alpha) \underbrace{\sum_{y=1}^{N} f(x, y) h_r(y, \beta)}_{f h_r \equiv s}$$

(Here, we consider f and h_r as matrices). Let $s \equiv f h_r$. Then:

$$g(\alpha, \beta) = \sum_{x=1}^{N} h_c(x, \alpha) s(x, \beta) = \sum_{x=1}^{N} h_c^T(\alpha, x) s(x, \beta).$$

$\therefore \boxed{g = h_c^T s = h_c^T f h_r}$

Definition 3.5. Suppose A and B are two matrices. The **Kronecker product** $A \otimes B$ is defined as:

$$A \otimes B = \begin{pmatrix} a_{11}B & \cdots & a_{1N}B \\ a_{21}B & \cdots & a_{2N}B \\ \vdots & & \vdots \\ a_{N1}B & \cdots & a_{NN}B \end{pmatrix},$$

where a_{ij} is the i-th row, j-th column entry of A.

Theorem 3.6. *Consider a separable linear image transformation, whose PSF is given by: $h(x, \alpha, y, \beta) = h_c(x, \alpha) h_r(y, \beta)$. The transformation matrix H is given by:*

$$H = h_r^T \otimes h_c^T.$$

Proof. Exercise.

4 Stacking operator

In image processing, a very important operator is called the stacking operator, which converts a 2D image to a column vector.

Definition 4.1. Define

$$V_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow \text{ row } n \text{ and } N_n \equiv \begin{bmatrix} \mathbf{0} \\ I_N \\ \mathbf{0} \end{bmatrix} \begin{matrix} \leftarrow (n-1)\ N \times N \text{ zero matrix} \\ \leftarrow N \times N \text{ identity matrix} \\ \leftarrow (N-n)\ N \times N \text{ zero matrix} \end{matrix}$$

Let $f \in \mathcal{I}$ ($N \times N$ image). We define the **stacking operator** on f as:

$$\mathcal{S}f \equiv \vec{f} \equiv \sum_{n=1}^{N} N_n f V_n.$$

Remark. *We can check that:*

1. $\mathcal{S}f \in \mathbb{R}^{N^2 \times 1}$;
2. *The* 1^{st} *column of* f *forms the first* N *elements of* $\mathcal{S}f$,
   the 2^{nd} *column of* f *forms the second* N *elements of* $\mathcal{S}f$, *etc.*
3. $\mathcal{S}$ *is important for actual MATLAB implementation.*

Theorem 4.2. $\mathcal{S}$ *is linear. Also* $f = \sum_{n=1}^{N} N_n^T \vec{f} V_n^T$.

The proof is left as exercise.

5 Similarity measure between images

In image processing, we often need to approximate an image by another image with better properties. For example, the main idea of image denoising (removing artifacts/noises from image) is to approximate an input noisy image by a 'smoother' image. In order to approximate an image, it is necessary to have a measure to quantify the similarity between different images.
Recall that a digital image can be considered as a matrix. To measure the similarity between two images, it is equivalent to defining a matrix norm. We first recall the definition of a vector norm.

Definition 5.1. A **vector norm** is a function $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ satisfying the following conditions:

1. $\|\mathbf{x}\| \geq 0$, $\|\mathbf{x}\| = 0$ only if $\mathbf{x} = \mathbf{0}$;
2. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ (triangle inequality);
3. $\|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\|$;

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$.

The most commonly used vector norms are the vector p-norms.

Definition 5.2. Let $\mathbf{v} = (v_1, v_2, \ldots, v_n)^T \in \mathbb{R}^n$, and let $p \geq 1$. The **vector p-norm** of $\mathbf{v}$, denoted by $\|\mathbf{v}\|_p$, is given by

$$\|\mathbf{v}\|_p := \left(\sum_{i=1}^{n} |v_i|^p \right)^{\frac{1}{p}}.$$

The limiting case for $p \to \infty$ is given by

$$\|\mathbf{v}\|_\infty := \lim_{p \to \infty} \|\mathbf{v}\|_p = \max_{1 \leq i \leq n} |v_i|,$$

and is also called the supremum norm of $\mathbf{v}$.

One can check with Definition 5.1 to verify that the p-norms are indeed vector norms.
Having defined vector norms, a matrix norm can be induced from each vector norm.

Definition 5.3. Let $A \in \mathbb{R}^{n\times m}$ and $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ be a vector norm. We define the **induced matrix norm** $\|A\|$ to be the smallest $C \in \mathbb{R}$ such that

$$\|A\mathbf{x}\| \leq \mathbf{C}\|\mathbf{x}\| \text{ for all } \mathbf{x} \in \mathbb{R}^{\mathbf{m}},$$

or equivalently,

$$\|A\| = \sup_{\mathbf{x}\in\mathbb{R}^m, \mathbf{x}\neq\mathbf{0}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} = \sup_{\mathbf{x}\in\mathbb{R}^m, \|\mathbf{x}\|=1} \|A\mathbf{x}\|.$$

However, not every matrix norm can be induced from a vector norm. In fact, matrix norms are defined in a similar manner to vector norms.

Definition 5.4. A **matrix norm** is a function $\|\cdot\| : \mathbb{R}^{n\times m} \to \mathbb{R}$ satisfying the following conditions:

1. $\|A\| \geq 0$, $\|A\| = 0$ only if $A = \mathbf{0}$;
2. $\|A + B\| \leq \|A\| + \|B\|$ (triangle inequality);
3. $\|\alpha A\| = |\alpha|\|A\|$;

for all $A, B \in \mathbb{R}^{n\times m}$ and $\alpha \in \mathbb{R}$.

For example, having defined the stacking operator and vector p-norms, another set of matrix p-norms can be defined as the vector p-norms of the stacked versions of matrices.

Definition 5.5. Let $A \in \mathbb{R}^{n\times m}$, and let $p \geq 1$. The **entrywise matrix p-norm** of A, denoted by $\|A\|_{p,e}$, is given by

$$\|A\|_{p,e} := \|\mathcal{S}A\|_p = \left(\sum_{i=1}^{n}\sum_{j=1}^{m} |A(i,j)|^p\right)^{\frac{1}{p}}.$$

$\|A\|_{2,e}$ is also called the Frobenius norm (F-norm) of A; it is also denoted by $\|A\|_F$. Let $\mathbf{a}_j$ be the j-th column of A. We have

$$\|A\|_F = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{m} A(i,j)^2} = \sqrt{\sum_{j=1}^{m} \|\mathbf{a}_j\|_2^2} = \sqrt{\operatorname{tr}(A^T A)},$$

where $\operatorname{tr}(\cdot)$ is the trace of the matrix in the argument.
The limiting case for $p \to \infty$ is given by

$$\|A\|_\infty := \lim_{p\to\infty} \|A\|_p = \max_{\substack{1\leq i\leq n \\ 1\leq j\leq m}} |A(i,j)|,$$

and is also called the entrywise supremum norm of A.

Remark. *In literature, the notations $\|A\|_p$ and $\|A\|_\infty$ are used for both the induced norms and the entrywise norms. Unless otherwise specified, in the notes we reserve these notations for induced norms, and denote the entrywise norms by $\|A\|_{p,e}$ and $\|A\|_{\infty,e}$.*

One may check with Definition 5.4 to verify that all induced norms and entrywise norms are indeed matrix norms.

Theorem 5.6. *The induced matrix 2-norm and the F-norm are invariant under multiplication by unitary matrices (a unitary matrix U satisfies $U^*U = UU^* = I$; its column vectors are orthonormal and its row vectors are also orthonormal), i.e. for any $A \in \mathbb{R}^{n\times m}$ and unitary $U \in \mathbb{R}^{n\times n}$, we have $\|UA\|_2 = \|A\|_2$ and $\|UA\|_F = \|A\|_F$.*

Proof. Since for any $x \in \mathbb{R}^m$,

$$\|UA\mathbf{x}\|_2^2 = (UA\mathbf{x})^T(UA\mathbf{x}) = \mathbf{x}^T A^T U^T UA\mathbf{x} = \mathbf{x}^T A^T A\mathbf{x} = \|A\mathbf{x}\|_2^2,$$

we have

$$\|UA\|_2 = \max_{\|\mathbf{x}\|_2=1} \|UA\mathbf{x}\|_2 = \max_{\|\mathbf{x}\|_2=1} \|A\mathbf{x}\|_2 = \|A\|_2.$$

Furthermore,

$$\|UA\|_F = \sqrt{\operatorname{tr}\big((UA)^T(UA)\big)} = \sqrt{\operatorname{tr}(A^T U^T UA)} = \sqrt{\operatorname{tr}(A^T A)} = \|A\|_F.$$

□

With matrix norms defined, we can measure the dissimilarity between two matrices (or images) by computing the norms of their difference matrix. Among the entrywise p-norms, the 1-norm and 2-norm are the most frequently used dissimilarity measures. The following figures demonstrate their different emphases.

Figure 1: The images on the left and on the right are equally similar to the image in the middle in terms of the entrywise 1-norm. On the other hand, the image on the right is significant less similar to the image in the middle in terms of the entrywise 2-norm than the image on the left.

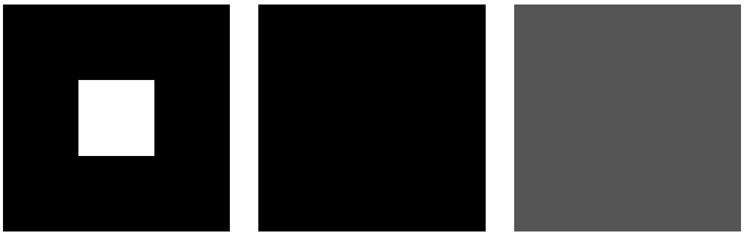

Figure 2: The images on the left and on the right are equally similar to the image in the middle in terms of the entrywise 2-norm. On the other hand, the image on the right is significant less similar to the image in the middle in terms of the entrywise 1-norm than the image on the left.

As seen from the figures, the 1-norm is more sensitive to widespread deviation in large regions, whereas the 2-norm is more sensitive to extreme pixel value differences, even if they are restricted to small regions. This trend goes on across different values of $p \geq 1$.

Exercises

1. In Example 2.7, what are $h(2,3,2,1)$ and $h(1,2,2,3)$?

2. Consider an image transformation on a 2×2 image. Suppose the matrix representation of the image transformation is given by:

$$H = \begin{pmatrix} 8 & 12 & 16 & 24 \\ 16 & 4 & 32 & 8 \\ 6 & 9 & 4 & 6 \\ 12 & 3 & 8 & 2 \end{pmatrix}.$$

Is the image transformation separable? Please explain in details. If yes, find g_1 and g_2 such that:

$$h(x, \alpha, y, \beta) = g_1(x, \alpha) g_2(y, \beta).$$

3. Suppose $H \in M_{4\times 4}$ is applied to a 2×2 image. Let

$$H = \begin{pmatrix} 5 & 3 & 2 & 8 \\ 3 & 2 & 1 & 1 \\ 3 & 1 & 1 & 1 \\ 2 & 2 & 5 & 3 \end{pmatrix}.$$

Is H shift-invariant? Please explain your answer with details.

4. Prove Theorem 3.4.

5. Prove Theorem 3.6.

6. Let $f = (f(m,n))_{-2\le m,n\le 2}$ be a 5×5 image. Consider a filter $H = (h(m,n))_{-2\le m,n\le 2}$, which is another 5×5 image. Note that the indices are taken from -2 to 2. Suppose $H = (a_1, a_2, a_3, a_4, a_5)^T(b_1, b_2, b_3, b_4, b_5)$.

   (a) Define the discrete convolution $H * f$.

   (b) Show that $H * f = H_1 * (H_2 * f)$, where

$$H_1 = \begin{pmatrix} 0 & 0 & a_1 & 0 & 0 \\ 0 & 0 & a_2 & 0 & 0 \\ 0 & 0 & a_3 & 0 & 0 \\ 0 & 0 & a_4 & 0 & 0 \\ 0 & 0 & a_5 & 0 & 0 \end{pmatrix} \text{ and } H_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ b_1 & b_2 & b_3 & b_4 & b_5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

   Hence, $H * f$ can be computed by a sequence of 1D convolutions.

7. Let H be a $(2N+1)\times(2N+1)$ matrix. Let $\mathcal{I}$ be the collection of $(2N+1)\times(2N+1)$ images. Assume the indices are taken from $-N$ to N. Define: $\mathcal{O}(H) : \mathcal{I} \to \mathcal{I}$ by:

$$\mathcal{O}(I) = I * H$$

   where $I * H$ refers to the discrete convolution.

   (a) Give the definition of discrete convolution. Argue that $\mathcal{O}$ is a linear operator and shift-invariant.

   (b) Show that $I * (H_1 * H_2) = (I * H_1) * H_2$, where H_1 and H_2 are $(2N+1)\times(2N+1)$ matrices.

   (c) Show that $I * H = H * I$.

8. For an $N\times N$ image g of real entries, let $g = UfV^T$, where U, V, f are $N\times N$ real matrices.

   (a) Show that

$$g = \sum_{i=1}^{N}\sum_{j=1}^{N} f_{ij}\vec{u}_i\vec{v}_j^T$$

   where

$$U = \begin{pmatrix} | & | & & | \\ \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_N \\ | & | & & | \end{pmatrix} \text{ and } V^T = \begin{pmatrix} \text{——} & \vec{v}_1^T & \text{——} \\ \text{——} & \vec{v}_2^T & \text{——} \\ & \vdots & \\ \text{——} & \vec{v}_N^T & \text{——} \end{pmatrix}$$

   (b) (Amended: Oct 18) Show that if f is diagonal, then the trace of g,

$$\mathrm{tr}(g) = \sum_{k=1}^{N} g_{kk} = \sum_{k=1}^{N}\sum_{l=1}^{N} f_{ll}u_{kl}v_{kl}.$$

9. Prove that $\mathcal{S}$ is linear and $f = \sum_{n=1}^{N} N_n^T\vec{f}V_n^T$.

10. For the following point-spread functions, determine whether they are (i) shift-invariant; (ii) separable. Prove your answer or provide a counterexample.

   (a) $h(x,\alpha,y,\beta) = \begin{cases} |(\alpha-x)(\beta-y)| & \text{if } |\alpha-x|\le 2, |\beta-y|\le 3 \\ 0 & \text{otherwise} \end{cases}$

   (b) $h(x,\alpha,y,\beta) = \sqrt{(\alpha-x)^4+(\beta-y)^3}$

   (c) $h(x,\alpha,y,\beta) = \begin{cases} \sqrt{17-(\alpha-x)^3+(\beta-y)^2} & \text{if } |\alpha-x|\le 2, |\beta-y|\le 3 \\ 0 & \text{otherwise} \end{cases}$

11. Let $H_1 = \begin{pmatrix} 4 & 5 & 7 & 3 \\ 3 & 4 & 5 & 7 \\ 7 & 3 & 4 & 5 \\ 5 & 7 & 3 & 4 \end{pmatrix}$ and $H_2 = \begin{pmatrix} 0 & 3 & 2 & 1 & 4 & 5 & 2 & 8 & 9 \\ 2 & 0 & 3 & 5 & 1 & 4 & 9 & 2 & 8 \\ 3 & 2 & 0 & 4 & 5 & 1 & 8 & 9 & 2 \\ 2 & 8 & 9 & 0 & 3 & 2 & 1 & 4 & 5 \\ 9 & 2 & 8 & 2 & 0 & 3 & 5 & 1 & 4 \\ 8 & 9 & 2 & 3 & 2 & 0 & 4 & 5 & 1 \\ 1 & 4 & 5 & 2 & 8 & 9 & 0 & 3 & 2 \\ 5 & 1 & 4 & 9 & 2 & 8 & 2 & 0 & 3 \\ 4 & 5 & 1 & 8 & 9 & 2 & 3 & 2 & 0 \end{pmatrix}$.

Discuss whether H_1 and H_2 represent shift-invariant linear transformations (with h_s being N-periodic in both arguments) on $N \times N$ square images. Please explain your answer with details.

12. Let $H_1 = \begin{pmatrix} 4 & 3 & 8 & 6 \\ 2 & 1 & 4 & 2 \\ 12 & 9 & 16 & 12 \\ 6 & 3 & 8 & 4 \end{pmatrix}$ and $H_2 = \begin{pmatrix} 1 & 2 & 1 & 2 \\ 3 & 4 & 3 & 4 \\ 5 & 6 & 5 & 6 \\ 7 & 8 & 7 & 8 \end{pmatrix}$.

Discuss whether H_1 and H_2 represent separable linear transformations on square images. Please explain your answer with details.

13. Let f and g be two $m \times n$ images. Assume that f and g are periodically extended.

    (a) Show that $f * g = g * f$, where $*$ denotes the convolution.

    (b) Let $f = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ and $g = \begin{pmatrix} 0 & 3 \\ 2 & 5 \end{pmatrix}$. Compute $f * g$.

14. Prove that $\|\cdot\|_p$ is a vector norm.

15. Show that all induced norms and entrywise norms are matrix norms.

16. (a) Let $A = \begin{pmatrix} 8 & 9 & 2 \\ 9 & 6 & 5 \\ 1 & 0 & 9 \end{pmatrix}$ and let $B = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.

    i. What is the value of α that minimizes $\|A - \alpha B\|_F$?

    ii. What is the value of α that minimizes $\|A - \alpha B\|_{1,e}$?

    (b) Let $C = \begin{pmatrix} 7 & 7 & 0 & 9 \\ 3 & 0 & 8 & 0 \end{pmatrix}$ and let $D = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$.

    i. What is the value of α that minimizes $\|C - \alpha D\|_F$?

    ii. What is the range of values of α that minimizes $\|C - \alpha D\|_{1,e}$?

    (c) Which central measures (mean, median, mode) of the pixel values are the values of α that respectively minimize the Frobenius norm difference and the entrywise 1-norm difference?

17. Consider the following two families of transformations:

    - $\mathcal{O}_1$: adding every pixel value by the same number, e.g. $\begin{pmatrix} 2 & 3 \\ 5 & 7 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 \\ 4 & 6 \end{pmatrix}$ by adding -1 to each pixel;
    - $\mathcal{O}_2$: scaling the difference of every entry with the mean pixel value, e.g. $\begin{pmatrix} 8 & 0 \\ 2 & 6 \end{pmatrix} \mapsto \begin{pmatrix} 6 & 2 \\ 3 & 5 \end{pmatrix}$ by halving the difference of every pixel value with the mean pixel value, which is 4.

    Let $A = \begin{pmatrix} 7 & 3 \\ 1 & 5 \end{pmatrix}$ and let $B = \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix}$.

    (a) Among all A_1 that can be obtained by transforming A via $\mathcal{O}_1$, what is the minimum value of $\|A_1 - B\|_F$?

(b) Among all A_2 that can be obtained by transforming A via $\mathcal{O}_2$, what is the minimum value of $\|A_2 - B\|_F$?

(c) Among all A_3 that can be obtained by transforming A via $\mathcal{O}_1$ and/or $\mathcal{O}_2$, what is the minimum value of $\|A_3 - B\|_F$?

Remark. *$\mathcal{O}_1$ and $\mathcal{O}_2$ correspond loosely to changing image brightness and contrast respectively.*

Appendix

Definition 3.7. A matrix $A \in M_{N\times N}(\mathbb{R})$ is said to be **Toeplitz** or **diagonal-constant** if $a_{ij} = a_{i+k,j+k}$ for any $i, j, k \in \mathbb{Z}$ such that $1 \leq i, j, i+k, j+k \leq N$. In other words,

$$A = \begin{pmatrix} d_0 & d_{-1} & d_{-2} & \cdots & d_{1-N} \\ d_1 & d_0 & d_{-1} & \cdots & d_{2-N} \\ d_2 & d_1 & d_0 & \cdots & d_{3-N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_{N-1} & d_{N-2} & d_{N-3} & \cdots & d_0 \end{pmatrix}$$

with $\{d_k : k = 1-N, 2-N, \ldots, N-1\} \subseteq \mathbb{R}$; the subscripts correspond to the values of $i - j$ on the particular (super-/sub-)diagonals.

Definition 3.8. A matrix $A \in M_{N^2\times N^2}(\mathbb{R})$ is said to be **block-Toeplitz** if it is of the following form:

$$A = \begin{pmatrix} A_0 & A_{-1} & A_{-2} & \cdots & A_{1-N} \\ A_1 & A_0 & A_{-1} & \cdots & A_{2-N} \\ A_2 & A_1 & A_0 & \cdots & A_{3-N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{N-1} & A_{N-2} & A_{N-3} & \cdots & A_0 \end{pmatrix},$$

where each block A_k is a $N \times N$ Toeplitz matrix.

Hence circulant matrices are Toeplitz, and block-circulant matrices are block-Toeplitz.
The proofs of Theorems 3.3 and 3.4 actually establish that linear transformations on square images with shift-invariant PSFs have block-Toeplitz transformation matrices. To establish that the matrices are block-circulant, it is necessary to assume that h_s is N-periodic in both arguments.