# MMAT5390: Mathematical Image Processing

## Solutions to Chapter 1 Exercises

1. $h(x,\alpha,y,\beta) = H(\alpha+(\beta-1)N, x+(y-1)N)$.
   $h(2,3,2,1) = H(3,5) = 0$, $h(1,2,2,3) = H(8,4) = 0$.

2. Note that
   $h(1,1,1,1) = 8$, $h(2,1,1,1) = 12$, $h(1,1,2,1) = 16$, $h(2,1,2,1) = 24$,
   $h(1,2,1,1) = 16$, $h(2,2,1,1) = 4$, $h(1,2,2,1) = 32$, $h(2,2,2,1) = 8$,
   $h(1,1,1,2) = 6$, $h(2,1,1,2) = 9$, $h(1,1,2,2) = 4$, $h(2,1,2,2) = 6$,
   $h(1,2,1,2) = 12$, $h(2,2,1,2) = 3$, $h(1,2,2,2) = 8$, $h(2,2,2,2) = 2$.
   Define $g_1 : \{1,2\}^2 \to \mathbb{R}$ by
   $g_1(1,1) = 2$, $g_1(2,1) = 3$, $g_1(1,2) = 4$, $g_1(2,2) = 1$
   and $g_2 : \{1,2\}^2 \to \mathbb{R}$ by
   $g_2(1,1) = 4$, $g_2(2,1) = 8$, $g_2(1,2) = 3$, $g_2(2,2) = 2$.
   As $h(x,\alpha,y,\beta) = g_1(x,\alpha)g_2(y,\beta)$ for all $1 \le x,\alpha,y,\beta) \le 2$, $h$ is separable.
   Alternatively, observe that
   $$H = \begin{pmatrix} 8 & 12 & 16 & 24 \\ 16 & 4 & 32 & 8 \\ 6 & 9 & 4 & 6 \\ 12 & 3 & 8 & 2 \end{pmatrix} = \begin{pmatrix} 4 & 8 \\ 3 & 2 \end{pmatrix} \otimes \begin{pmatrix} 2 & 3 \\ 4 & 1 \end{pmatrix}.$$
   Hence $h$ is separable.

3. Suppose there exists $g : \{-1,0,1\}^2 \to \mathbb{R}$ such that $h(x,\alpha,y,\beta) = g(\alpha-x,\beta-y)$ for all $1 \le x,\alpha,y,\beta \le 2$.
   Then $3 = h(2,1,1,1) = g(-1,0) = h(2,1,2,2) = 1$. Contradiction.
   Hence $h$ is not shift-invariant.
   Alternatively, observe that $\begin{pmatrix} 2 & 8 \\ 1 & 1 \end{pmatrix}$ is not Toeplitz; thus $h$ is not shift-invariant.

4. Let $h$ be the shift-invariant PSF of a linear image transformation on $M_{n\times n}(\mathbb{R})$, with $h_s$ $n$-periodic in both arguments such that $h(x,\alpha,y,\beta) = h_s(\alpha-x,\beta-y)$. Let $H$ be the corresponding transformation matrix.
   Let $a \in \mathbb{Z}$. Then
   $$\begin{aligned} H\left(\alpha+(\beta+a-1)n, x+(y+a-1)n\right) &= h(x,\alpha,y+a,\beta+a) \\ &= h_s\left(\alpha-x,(\beta+a)-(y+a)\right) \\ &= h_s(\alpha-x,\beta-y) \\ &= h(x,\alpha,y,\beta) \\ &= H\left(\alpha+(\beta-1)n, x+(y-1)n\right). \end{aligned}$$
   Also, by periodicity of $h_s$, for $y \in \mathbb{N} \cap [1,n-1]$,
   $$\begin{aligned} H(\alpha+(n-1)n, x+(y-1)n) &= h(x,\alpha,y,n) \\ &= h_s(\alpha-x,n-y) \\ &= h_s(\alpha-x,1-(y+1) \\ &= h(x,\alpha,y+1,1) \\ &= H(\alpha,x+yn) \end{aligned}$$

and for $\beta \in \mathbb{N} \cap [1, n-1]$,

$$\begin{aligned} H(\alpha + (\beta - 1)n, x + (n-1)n) &= h(x, \alpha, n, \beta) \\ &= h_s(\alpha - x, \beta - n) \\ &= h_s(\alpha - x, (\beta + 1) - 1) \\ &= h(x, \alpha, 1, \beta + 1) \\ &= H(\alpha + \beta n, x). \end{aligned}$$

Hence $H$ is circulant when viewed as a matrix consisting of blocks of fixed $(y, \beta)$-values. Combined with the result of Theorem 1.13, we establish that $H$ is block-circulant.

5. Let $h$ be the separable PSF of a linear image transformation, with $h(x, \alpha, y, \beta) = h_c(x, \alpha) h_r(y, \beta)$. Let $H$ be the corresponding transformation matrix.

   Then the $y = k, \beta = l$-submatrix of $H$ (denoted by $\tilde{H}_{kl}$) is given by

$$\begin{aligned} \begin{pmatrix} & x \rightarrow \\ \alpha \downarrow & \begin{pmatrix} y = k \\ \beta = l \end{pmatrix} \end{pmatrix} &= [H(\alpha + (l-1)n, x + (k-1)n)]_{\substack{1 \le x \le n \\ 1 \le \alpha \le n}} \\ &= [h(x, \alpha, k, l)]_{\substack{1 \le x \le n \\ 1 \le \alpha \le n}} \\ &= [h_c(x, \alpha) h_r(k, l)]_{\substack{1 \le x \le n \\ 1 \le \alpha \le n}} \\ &= h_r(k, l)[h_c(x, \alpha)]_{\substack{1 \le x \le n \\ 1 \le \alpha \le n}} \\ &= h_r(k, l) h_c^T. \end{aligned}$$

   Recall that

$$H = \begin{pmatrix} \begin{pmatrix} & x \rightarrow \\ \alpha \downarrow & \begin{pmatrix} y = 1 \\ \beta = 1 \end{pmatrix} \end{pmatrix} & \begin{pmatrix} & x \rightarrow \\ \alpha \downarrow & \begin{pmatrix} y = 2 \\ \beta = 1 \end{pmatrix} \end{pmatrix} & \cdots & \begin{pmatrix} & x \rightarrow \\ \alpha \downarrow & \begin{pmatrix} y = n \\ \beta = 1 \end{pmatrix} \end{pmatrix} \\ \begin{pmatrix} & x \rightarrow \\ \alpha \downarrow & \begin{pmatrix} y = 1 \\ \beta = 2 \end{pmatrix} \end{pmatrix} & \begin{pmatrix} & x \rightarrow \\ \alpha \downarrow & \begin{pmatrix} y = 2 \\ \beta = 2 \end{pmatrix} \end{pmatrix} & \cdots & \begin{pmatrix} & x \rightarrow \\ \alpha \downarrow & \begin{pmatrix} y = n \\ \beta = 2 \end{pmatrix} \end{pmatrix} \\ \vdots & \vdots & \ddots & \vdots \\ \begin{pmatrix} & x \rightarrow \\ \alpha \downarrow & \begin{pmatrix} y = 1 \\ \beta = n \end{pmatrix} \end{pmatrix} & \begin{pmatrix} & x \rightarrow \\ \alpha \downarrow & \begin{pmatrix} y = 2 \\ \beta = n \end{pmatrix} \end{pmatrix} & \cdots & \begin{pmatrix} & x \rightarrow \\ \alpha \downarrow & \begin{pmatrix} y = n \\ \beta = n \end{pmatrix} \end{pmatrix} \end{pmatrix}$$

$$= \begin{pmatrix} \tilde{H}_{11} & \tilde{H}_{21} & \cdots & \tilde{H}_{n1} \\ \tilde{H}_{12} & \tilde{H}_{22} & \cdots & \tilde{H}_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{H}_{1n} & \tilde{H}_{2n} & \cdots & \tilde{H}_{nn} \end{pmatrix} = \begin{pmatrix} h_r(1,1)h_c^T & h_r(2,1)h_c^T & \cdots & h_r(n,1)h_c^T \\ h_r(1,2)h_c^T & h_r(2,2)h_c^T & \cdots & h_r(n,2)h_c^T \\ \vdots & \vdots & \ddots & \vdots \\ h_r(1,n)h_c^T & h_r(2,n)h_c^T & \cdots & h_r(n,n)h_c^T \end{pmatrix}$$

$$= \begin{pmatrix} h_r^T(1,1)h_c^T & h_r^T(1,2)h_c^T & \cdots & h_r^T(1,n)h_c^T \\ h_r^T(2,1)h_c^T & h_r^T(2,2)h_c^T & \cdots & h_r^T(2,n)h_c^T \\ \vdots & \vdots & \ddots & \vdots \\ h_r^T(n,1)h_c^T & h_r^T(n,2)h_c^T & \cdots & h_r^T(n,n)h_c^T \end{pmatrix} = h_r^T \otimes h_c^T.$$

6. Assume $f$ and $H$ are periodically extended.

   (a) $H * f$ is the $5 \times 5$ matrix whose entries are given by

$$H * f(\alpha, \beta) = \sum_{m=-2}^{2} \sum_{n=-2}^{2} H(m, n) f(\alpha - m, \beta - n).$$

(b)

$$
\begin{aligned}
H * f(\alpha,\beta) &= \sum_{m=-2}^{2}\sum_{n=-2}^{2} H(m,n) f(\alpha-m,\beta-n) \\
&= \sum_{m=-2}^{2}\sum_{n=-2}^{2} a_{m+3} b_{n+3} f(\alpha-m,\beta-n) \\
&= \sum_{m=-2}^{2}\sum_{n=-2}^{2} H_1(m,0) H_2(0,n) f(\alpha-m,\beta-n) \\
&= \sum_{n=-2}^{2} H_2(0,n) \sum_{m=-2}^{2}\sum_{n'=-2}^{2} H_1(m,n') f(\alpha-m,\beta-n-n') \\
&= \sum_{n=-2}^{2} H_2(0,n) H_1 * f(\alpha,\beta-n) \\
&= \sum_{m'=-2}^{2}\sum_{n=-2}^{2} H_2(m',n) H_1 * f(\alpha-m',\beta-n) \\
&= H_2 * (H_1 * f)(\alpha,\beta).
\end{aligned}
$$

Hence $H * f = H_1 * (H_2 * f)$.

7. (a) Assume $I_1, I_2 \in \mathcal{I}$ are periodically extended.
The discrete convolution $I_1 * I_2$ of $I_1$ and $I_2$ is the $(2N+1)\times(2N+1)$ matrix whose entries are defined by

$$
I_1 * I_2(\alpha,\beta) = \sum_{m=-N}^{N}\sum_{n=-N}^{N} I_1(m,n) I_2(\alpha-m,\beta-n).
$$

Let $I_1, I_2 \in \mathcal{I}$, and let $c \in \mathbb{R}$. Then

$$
\begin{aligned}
\mathcal{O}(I_1+cI_2) &= (I_1+cI_2) * H \\
&= \left[\sum_{m=-N}^{N}\sum_{n=-N}^{N} (I_1+cI_2)(m,n) H(\alpha-m,\beta-n)\right]_{-N\le\alpha,\beta\le N} \\
&= \left[\sum_{m=-N}^{N}\sum_{n=-N}^{N} [I_1(m,n)H(\alpha-m,\beta-n) + cI_2(m,n)H(\alpha-m,\beta-n)]\right]_{-N\le\alpha,\beta\le N} \\
&= [I_1 * H(\alpha,\beta) + cI_2 * H(\alpha,\beta)]_{-N\le\alpha,\beta\le N} \\
&= I_1 * H + cI_2 * H \\
&= \mathcal{O}(I_1) + c\mathcal{O}(I_2).
\end{aligned}
$$

Hence $\mathcal{O}$ is linear.
For any $I \in \mathcal{I}$, the PSF $h$ of $\mathcal{O}$ satisfies

$$
\sum_{x=-N}^{N}\sum_{y=-N}^{N} h(x,\alpha,y,\beta) I(x,y) = I * H(\alpha,\beta) = \sum_{m=-N}^{N}\sum_{n=-N}^{N} I(m,n) H(\alpha-m,\beta-n),
$$

and thus $h(x,\alpha,y,\beta) = H(\alpha-x,\beta-y)$.
Hence $h$ is shift-invariant.

(b) Let $H_1, H_2 \in \mathcal{I}$. Then

$$
\begin{aligned}
I * (H_1 * H_2)(\alpha,\beta) &= \sum_{m=-N}^{N}\sum_{n=-N}^{N} I(m,n) H_1 * H_2(\alpha-m,\beta-n) \\
&= \sum_{m=-N}^{N}\sum_{n=-N}^{N} I(m,n) \sum_{m'=-N}^{N}\sum_{n'=-N}^{N} H_1(m',n') H_2(\alpha-m-m',\beta-n-n')
\end{aligned}
$$

$$= \sum_{m=-N}^{N} \sum_{n=-N}^{N} I(m,n) \sum_{m''=-N+m}^{N+m} \sum_{n''=-N+n}^{N+n} H_1(m''-m, n''-n) H_2(\alpha - m'', \beta - n'')$$

$$= \begin{cases} \sum\limits_{m=-N}^{N} \sum\limits_{n=-N}^{N} I(m,n) \left( \sum\limits_{m''=-N+m}^{-N-1} + \sum\limits_{m''=-N}^{N+m} \right) \sum\limits_{n''=-N+n}^{N+n} \\ \quad H_1(m''-m, n''-n) H_2(\alpha - m'', \beta - n'') \text{ if } -N \leq m \leq 0 \\ \sum\limits_{m=-N}^{N} \sum\limits_{n=-N}^{N} I(m,n) \left( \sum\limits_{m''=-N+m}^{N} + \sum\limits_{m''=N+1}^{N+m} \right) \sum\limits_{n''=-N+n}^{N+n} \\ \quad H_1(m''-m, n''-n) H_2(\alpha - m'', \beta - n'') \text{ if } 1 \leq m \leq N \end{cases}$$

$$= \sum_{m=-N}^{N} \sum_{n=-N}^{N} I(m,n) \sum_{m''=-N}^{N} \sum_{n''=-N+n}^{N+n} H_1(m''-m, n''-n) H_2(\alpha - m'', \beta - n'')$$

$$= \begin{cases} \sum\limits_{m=-N}^{N} \sum\limits_{n=-N}^{N} I(m,n) \sum\limits_{m''=-N}^{N} \left( \sum\limits_{n''=-N+n}^{-N-1} + \sum\limits_{n''=-N}^{N+n} \right) \\ \quad H_1(m''-m, n''-n) H_2(\alpha - m'', \beta - n'') \text{ if } -N \leq n \leq 0 \\ \sum\limits_{m=-N}^{N} \sum\limits_{n=-N}^{N} I(m,n) \sum\limits_{m''=-N}^{N} \left( \sum\limits_{n''=-N+n}^{N} + \sum\limits_{n''=N+1}^{N+n} \right) \\ \quad H_1(m''-m, n''-n) H_2(\alpha - m'', \beta - n'') \text{ if } 1 \leq n \leq N \end{cases}$$

$$= \sum_{m=-N}^{N} \sum_{n=-N}^{N} I(m,n) \sum_{m''=-N}^{N} \sum_{n''=-N}^{N} H_1(m''-m, n''-n) H_2(\alpha - m'', \beta - n'')$$

$$= \sum_{m''=-N}^{N} \sum_{n''=-N}^{N} I * H_1(m'', n'') H_2(\alpha - m'', \beta - n'')$$

$$= (I * H_1) * H_2(\alpha, \beta).$$

Hence $I * (H_1 * H_2) = (I * H_1) * H_2$.

(c)

$$\begin{aligned} I * H(\alpha, \beta) &= \sum_{m=-N}^{N} \sum_{n=-N}^{N} I(m,n) H(\alpha - m, \beta - n) \\ &= \sum_{m'=\alpha-N}^{\alpha+N} \sum_{n'=\beta-N}^{\beta+N} I(\alpha - m', \beta - n') H(m', n') \\ &= \begin{cases} \left( \sum\limits_{m'=\alpha-N}^{-N-1} + \sum\limits_{m'=-N}^{\alpha+N} \right) \sum\limits_{n'=\beta-N}^{\beta+N} I(\alpha - m', \beta - n') H(m', n') & \text{if } -N \leq \alpha \leq 0 \\ \left( \sum\limits_{m'=\alpha-N}^{N} + \sum\limits_{m'=N+1}^{\alpha+N} \right) \sum\limits_{n'=\beta-N}^{\beta+N} I(\alpha - m', \beta - n') H(m', n') & \text{if } 1 \leq \alpha \leq N \end{cases} \\ &= \sum_{m'=-N}^{N} \sum_{n'=\beta-N}^{\beta+N} H(m', n') I(\alpha - m', \beta - n') \end{aligned}$$

$$= \begin{cases} \sum\limits_{m'=-N}^{N} \left( \sum\limits_{n'=\beta-N}^{-N-1} + \sum\limits_{n'=-N}^{\beta+N} \right) H(m', n') I(\alpha - m', \beta - n') & \text{if } -N \leq \beta \leq 0 \\ \sum\limits_{m'=-N}^{N} \left( \sum\limits_{n'=\beta-N}^{N} + \sum\limits_{n'=N+1}^{\beta+N} \right) H(m', n') I(\alpha - m', \beta - n') & \text{if } 1 \leq \beta \leq N \end{cases}$$

$$= \sum_{m'=-N}^{N} \sum_{n'=-N}^{N} H(m', n') I(\alpha - m', \beta - n')$$

$$= H * I(\alpha, \beta).$$

Hence $I * H = H * I$.

8. (a)

$$\begin{aligned} g_{kl} &= (UfV^T)_{kl} \\ &= \sum_{i=1}^{N} u_{ki}(fV^T)_{il} \\ &= \sum_{i=1}^{N} u_{ki} \sum_{j=1}^{N} f_{ij} v_{lj}. \end{aligned}$$

On the other hand,

$$\begin{aligned} (\sum_{i=1}^{N}\sum_{j=1}^{N} f_{ij}\vec{u}_i\vec{v}_j^T)_{kl} &= \sum_{i=1}^{N}\sum_{j=1}^{N} f_{ij}(\vec{u}_i\vec{v}_j^T)_{kl} \\ &= \sum_{i=1}^{N}\sum_{j=1}^{N} f_{ij}u_{ki}v_{lj}. \end{aligned}$$

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

(b) Suppose $f$ is diagonal. Then

$$\begin{aligned} \text{tr}(g) &= \sum_{k=1}^{N} g_{kk} \\ &= \sum_{k=1}^{N}\sum_{i=1}^{N}\sum_{j=1}^{N} f_{ij}u_{ki}v_{kj} \\ &= \sum_{k=1}^{N}\sum_{l=1}^{N} f_{ll}u_{kl}v_{kl}. \end{aligned}$$

9. Let $f, g \in M_{N\times N}(\mathbb{R})$ and $c \in \mathbb{R}$. Then

$$\begin{aligned} \mathcal{S}(f+cg) &= \sum_{n=1}^{N} N_n(f+cg)V_n \\ &= \sum_{n=1}^{N} N_n f V_n + c\sum_{n=1}^{N} N_n g V_n \\ &= \mathcal{S}f + c\mathcal{S}g. \end{aligned}$$

Hence $\mathcal{S}$ is linear.

In addition,

$$\begin{aligned} \sum_{n=1}^{N} N_n^T \vec{f} V_n^T &= \sum_{n=1}^{N}\sum_{m=1}^{N} N_n^T N_m f V_m V_n^T \\ &= \sum_{n=1}^{N}\sum_{m=1}^{N} \delta(n-m) I_N f E_{m,n} \\ &= f\sum_{n=1}^{N} E_{n,n} = f, \end{aligned}$$

where $E_{m,n}$ is the $N \times N$ matrix with all entries zero except $E_{m,n}(m,n) = 1$.

10. (a) i. Let $h_s(u,v) = \begin{cases} |uv| & \text{if } |u| \le 2, |v| \le 3 \\ 0 & \text{otherwise} \end{cases}$.

Then $h(x,\alpha,y,\beta) = h_s(\alpha - x, \beta - y)$.

Hence $h$ is shift-invariant.

ii. Let $h_r(x,\alpha) = \begin{cases} |\alpha - x| & \text{if } |\alpha - x| \le 2 \\ 0 & \text{otherwise} \end{cases}$ and $h_c(y,\beta) = \begin{cases} |\beta - y| & \text{if } |\beta - y| \le 3 \\ 0 & \text{otherwise} \end{cases}$.
Then $h(x,\alpha,y,\beta) = h_r(x,\alpha)h_c(y,\beta)$.
Hence $h$ is separable.

(b) i. Let $h_s(u,v) = \sqrt{u^4 + v^3}$.
Then $h(x,\alpha,y,\beta) = h_s(\alpha - x, \beta - y)$.
Hence $h$ is shift-invariant.

ii. Suppose $h$ is separable, i.e. there exists functions $h_r$ and $h_c$ such that $h(x,\alpha,y,\beta) = h_r(x,\alpha)h_c(y,\beta)$.
Then $h_r(0,0)h_c(0,0) = h(0,0,0,0) = 0$, which implies $h_r(0,0) = 0$ or $h_c(0,0) = 0$.
On the other hand, $h_r(0,0)h_c(1,0) = h(0,0,1,0) = 1$ and $h_r(1,0)h_c(0,0) = h(1,0,0,0) = 1$, which imply $h_r(0,0) \ne 0$ and $h_c(0,0) \ne 0$. Contradiction.
Hence $h$ is not separable.

(c) i. Let $h_s(u,v) = \begin{cases} \sqrt{17 - u^3 - v^2} & \text{if } |u| \le 2, |v| \le 3 \\ 0 & \text{otherwise} \end{cases}$.
Then $h(x,\alpha,y,\beta) = h_s(\alpha - x, \beta - y)$.
Hence $h$ is shift-invariant.

ii. Suppose $h$ is separable, i.e. there exist functions $h_r$ and $h_c$ such that $h(x,\alpha,y,\beta) = h_r(x,\alpha)h_c(y,\beta)$.
Then $h_r(0,0)h_c(0,0) = h(0,0,0,0) = \sqrt{17}$,
$h_r(1,0)h_c(0,0) = h(1,0,0,0) = 4$,
$h_r(0,0)h_c(1,0) = h(0,0,1,0) = 4$ and
$h_r(1,0)h_c(1,0) = h(1,0,1,0) = \sqrt{15}$.
But then

$$\begin{aligned} 16 &= [h_r(1,0)h_c(0,0)][h_r(0,0)h_c(1,0)] \\ &= [h_r(0,0)h_c(0,0)][h_r(1,0)h_c(1,0)] \\ &= \sqrt{255}. \end{aligned}$$

Contradiction.
Hence $h$ is not separable.

11. (a) Note that $\begin{pmatrix} 4 & 5 \\ 3 & 4 \end{pmatrix}$ and $\begin{pmatrix} 7 & 3 \\ 5 & 7 \end{pmatrix}$ are Toeplitz, and that $H_1$ is circulant (hence Toeplitz) when viewed as a matrix of $2 \times 2$ blocks.
Hence $H_1$ is block-Toeplitz, and thus represents a shift-invariant linear transformation on $2 \times 2$ images.
On the other hand, as $\begin{pmatrix} 4 & 5 \\ 3 & 4 \end{pmatrix}$ is not circulant, $H_1$ is not block-circulant. Hence $h_s$ is not 2-periodic in some of its arguments.

(b) Note that $\begin{pmatrix} 0 & 3 & 2 \\ 2 & 0 & 3 \\ 3 & 2 & 0 \end{pmatrix}$, $\begin{pmatrix} 1 & 4 & 5 \\ 5 & 1 & 4 \\ 4 & 5 & 1 \end{pmatrix}$ and $\begin{pmatrix} 2 & 8 & 9 \\ 9 & 2 & 8 \\ 8 & 9 & 2 \end{pmatrix}$ are circulant, and that $H_2$ is circulant when viewed as a matrix of $3 \times 3$ blocks.
Hence $H_2$ is block-circulant, and thus represents a shift-invariant linear transformation on $3 \times 3$ images with $h_s$ being 3-periodic in both arguments.

12. (a) Note that

$$H_1 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \otimes \begin{pmatrix} 4 & 3 \\ 2 & 1 \end{pmatrix}.$$

Hence $H_1$ represents a separable linear transformation on $2 \times 2$ images.

(b) Suppose $H_2$ represents a separable linear transformation on $3 \times 3$ images. Then there exists $A, B \in M_{3\times3}$ such that $H_2 = A \otimes B$.
Then $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = a_{11}B$ whereas $\begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = a_{21}B$.
But then $6 = (a_{11}b_{11})(a_{21}b_{12}) = (a_{11}b_{12})(a_{21}b_{11}) = 10$. Contradiction.
Hence $H_2$ does not represent a separable linear transformation on $3 \times 3$ images.

13. (a) Let $1 \le \alpha \le m$ and $1 \le \beta \le n$. Then

$$\begin{aligned}
f * g(\alpha, \beta) &= \sum_{i=1}^{m} \sum_{j=1}^{n} f(i,j) g(\alpha - i, \beta - j) \\
&= \sum_{i'=\alpha-m}^{\alpha-1} \sum_{j'=\beta-n}^{\beta-1} f(\alpha - i', \beta - j') g(i', j') \\
&= \left( \sum_{i'=\alpha-m}^{0} + \sum_{i'=1}^{\alpha-1} \right) \left( \sum_{j'=\beta-n}^{0} + \sum_{j'=1}^{\beta-1} \right) g(i', j') f(\alpha - i', \beta - j') \\
&= \sum_{i'=1}^{m} \sum_{j'=1}^{n} g(i', j') f(\alpha - i', \beta - j') \\
&= g * f(\alpha, \beta).
\end{aligned}$$

Hence $f * g = g * f$.

(b) For simplicity (and to guide the indexing of $f * g$), we only consider the cases where $f$ and $g$ are indexed with the same set of indices.

If $f$ and $g$ are indexed with $1 \le i, j \le 2$, i.e. if

$$f = (f(i,j))_{1 \le i,j \le 2} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \text{ and } g = (g(i,j))_{1 \le i,j \le 2} = \begin{pmatrix} 0 & 3 \\ 2 & 5 \end{pmatrix},$$

then

$$\begin{aligned}
f * g(1,1) &= f(1,1)g(2,2) + f(1,2)g(2,1) + f(2,1)g(1,2) + f(2,2)g(1,1) \\
&= 1 \cdot 5 + 2 \cdot 2 + 3 \cdot 3 + 4 \cdot 0 = 18, \\
f * g(1,2) &= f(1,1)g(2,1) + f(1,2)g(2,2) + f(2,1)g(1,1) + f(2,2)g(1,2) \\
&= 1 \cdot 2 + 2 \cdot 5 + 3 \cdot 0 + 4 \cdot 3 = 24, \\
f * g(2,1) &= f(1,1)g(1,2) + f(1,2)g(1,1) + f(2,1)g(2,2) + f(2,2)g(2,1) \\
&= 1 \cdot 3 + 2 \cdot 0 + 3 \cdot 5 + 4 \cdot 2 = 26, \text{ and} \\
f * g(2,2) &= f(1,1)g(1,1) + f(1,2)g(1,2) + f(2,1)g(2,1) + f(2,2)g(2,2) \\
&= 1 \cdot 0 + 2 \cdot 3 + 3 \cdot 2 + 4 \cdot 5 = 32,
\end{aligned}$$

i.e. $f * g = (f * g(i,j))_{1 \le i,j \le 2} = \begin{pmatrix} 18 & 24 \\ 26 & 32 \end{pmatrix}$.

If $f$ and $g$ are indexed with $0 \le i, j \le 1$, i.e. if

$$f = (f(i,j))_{0 \le i,j \le 1} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \text{ and } g = (g(i,j))_{0 \le i,j \le 1} = \begin{pmatrix} 0 & 3 \\ 2 & 5 \end{pmatrix},$$

then

$$\begin{aligned}
f * g(0,0) &= f(0,0)g(0,0) + f(0,1)g(0,1) + f(1,0)g(1,0) + f(1,1)g(1,1) \\
&= 1 \cdot 0 + 2 \cdot 3 + 3 \cdot 2 + 4 \cdot 5 = 32, \\
f * g(0,1) &= f(0,0)g(0,1) + f(0,1)g(0,0) + f(1,0)g(1,1) + f(1,1)g(1,0) \\
&= 1 \cdot 3 + 2 \cdot 0 + 3 \cdot 5 + 4 \cdot 2 = 26, \\
f * g(1,0) &= f(0,0)g(1,0) + f(0,1)g(1,1) + f(1,0)g(0,0) + f(1,1)g(0,1) \\
&= 1 \cdot 2 + 2 \cdot 5 + 3 \cdot 0 + 4 \cdot 3 = 24, \text{ and} \\
f * g(1,1) &= f(0,0)g(1,1) + f(0,1)g(1,0) + f(1,0)g(0,1) + f(1,1)g(0,0) \\
&= 1 \cdot 5 + 2 \cdot 2 + 3 \cdot 3 + 4 \cdot 0 = 18,
\end{aligned}$$

i.e. $f * g = (f * g(i,j))_{0 \le i,j \le 1} = \begin{pmatrix} 32 & 26 \\ 24 & 18 \end{pmatrix}$, which is the same as the previous example upon periodic extension.

One may verify that the result is the same when the indexing is

$$0 \le i \le 1 \text{ and } 1 \le j \le 2$$

or

$$1 \le i \le 2 \text{ and } 0 \le j \le 1$$

and confirm that $f * g$ takes the same values regardless of indexing, since all other index sets are periodic shifts from the above four.

**Remark**. In this example, since $f, g \in M_{2\times 2}$, $f * g$ does not depend on the indexing of $f$ and $g$. In general, however, when $f$ and $g$ are indexed with different sets of indices, results may differ in the sense that they are not the same even with periodic extension. A simple example is $f = g = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, which gives $f * g = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ when indexed with $0 \le i, j \le 2$ and $f * g = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ when indexed with $1 \le i, j \le 3$. Even after periodic extension, these results differ by a shift.

14. $\|\cdot\|_p : \mathbb{R}^n \to \mathbb{R}$ is defined by

$$\|v\|_p = \left( \sum_{i=1}^{n} |v_i|^p \right)^{\frac{1}{p}} \quad \text{for any } \vec{v} = (v_1, \cdots, v_n)^T \in \mathbb{R}^n.$$

(a) For any $\vec{v} = (v_1, \cdots, v_n)^T \in \mathbb{R}^n$,

$$\|v\|_p = \left( \sum_{i=1}^{n} |v_i|^p \right)^{\frac{1}{p}} \ge 0.$$

Since $\vec{0} = (\underbrace{0, \cdots, 0}_{n \text{ zeros}})^T$,

$$\|\vec{0}\|_p = \left( \sum_{i=1}^{n} 0^p \right)^{\frac{1}{p}} = 0.$$

On the other hand, for any $\vec{v} = (v_1, \cdots, v_n) \in \mathbb{R}^n$ such that $\|v\|_p = 0$,

$$\left( \sum_{i=1}^{n} |v_i|^p \right)^{\frac{1}{p}} = 0,$$

$$\sum_{i=1}^{n} |v_i|^p = 0,$$

$$|v_i|^p = 0 \text{ for each } i = 1, \cdots, n$$

and thus $\vec{v} = \vec{0}$.

Hence $\|v\|_p = 0$ if and only if $\vec{v} = \vec{0}$.

(b) This triangle inequality of $p$-norms is precisely Minkowski's inequality. For self-containment, we will establish it from scratch.

Let $\vec{v} = (v_1, \cdots, v_n)^T \in \mathbb{R}^n$ and $\vec{w} = (w_1, \cdots, w_n)^T \in \mathbb{R}^n$, and let $1 \le i \le n$.

For $p = 1$,

$$\|\vec{v} + \vec{w}\|_1 = \sum_{i=1}^{N} |v_i + w_i| \le \sum_{i=1}^{N} (|v_i| + |w_i|) = \sum_{i=1}^{N} |v_i| + \sum_{i=1}^{N} |w_i| = \|\vec{v}\|_1 + \|\vec{w}\|_1.$$

For $p > 1$, we will require Hölder's inequality, which is in turn derived from Young's inequality.

**Young's inequality**

Let $a, b \geq 0$, and let $p, q > 0$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Then

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Proof

WLOG assume $a > 0$ and $b > 0$.

Notice that $f(t) = \ln(t)$ is concave on the positive real axis:

$$f''(t) = -\frac{1}{t^2} < 0 \text{ for any } t \in (0, +\infty).$$

Then for any $t \in [0, 1]$,

$$t \ln(a^p) + (1-t)\ln(b^q) \leq \ln(ta^p + (1-t)b^q). \tag{1}$$

The constraint $\frac{1}{p} + \frac{1}{q} = 1$ restricts $p, q \geq 1$.

Hence by substituting $t = \frac{1}{p}$ into (1),

$$\ln(a) + \ln(b) \leq \ln(\frac{a^p}{p} + \frac{b^q}{q}),$$

and the result follows from exponentiating both sides.

**Hölder's inequality**

Let $\vec{v} = (v_1, \cdots, v_n)^T, \vec{w} = (w_1, \cdots, w_n)^T \in \mathbb{R}^n$, and let $p, q > 0$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Then

$$\sum_{i=1}^{n} v_i w_i \leq \left(\sum_{i=1}^{n} |v_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} |w_i|^q\right)^{\frac{1}{q}} = \|\vec{v}\|_p \|\vec{w}\|_q.$$

Proof

By (a), if $\|\vec{v}\|_p = 0$, then $\vec{v} = 0$ and $\sum\limits_{i=1}^{n} v_i w_i = 0$.

WLOG assume $\|\vec{v}\|_p \neq 0$ and $\|\vec{w}\|_q \neq 0$.

Then $\|\frac{\vec{v}}{\|\vec{v}\|_p}\|_p = \|\frac{\vec{w}}{\|\vec{w}\|_q}\|_q = 1$.

Then by Young's inequality,

$$\frac{v_i}{\|\vec{v}\|_p} \frac{w_i}{\|\vec{w}\|_q} \leq \frac{|v_i|}{\|\vec{v}\|_p} \frac{|w_i|}{\|\vec{w}\|_q} \leq \frac{|v_i|^p}{p\|\vec{v}\|_p^p} + \frac{|w_i|^q}{q\|\vec{w}\|_q^q},$$

and thus

$$\begin{aligned}\sum_{i=1}^{n} v_i w_i &\leq \|\vec{v}\|_p \|\vec{w}\|_q \sum_{i=1}^{n} \left(\frac{|v_i|^p}{p\|\vec{v}\|_p^p} + \frac{|w_i|^q}{q\|\vec{w}\|_q^q}\right) \\ &= \|\vec{v}\|_p \|\vec{w}\|_q \left(\frac{1}{p\|\vec{v}\|_p^p} \sum_{i=1}^{n} |v_i|^p + \frac{1}{q\|\vec{w}\|_q^q} \sum_{i=1}^{n} |w_i|^q\right) \\ &= \|\vec{v}\|_p \|\vec{w}\|_q (\frac{1}{p} + \frac{1}{q}) = \|\vec{v}\|_p \|\vec{w}\|_q.\end{aligned}$$

**Minkowski's inequality**

Let $\vec{v} = (v_1, \cdots, v_n)^T, \vec{w} = (w_1, \cdots, w_n)^T \in \mathbb{R}^n$, and let $p \geq 1$.

Then

$$\|\vec{v} + \vec{w}\|_p \leq \|\vec{v}\|_p + \|\vec{w}\|_p.$$

Proof

As we have covered $p = 1$, suppose $p > 1$.

Let $q > 0$ such that $\frac{1}{p} + \frac{1}{q} = 1$.
Note that

$$\begin{aligned} |v_i + w_i|^p &= |v_i + w_i|\,|v_i + w_i|^{p-1} \\ &\leq |v_i|\,|v_i + w_i|^{p-1} + |w_i|\,|v_i + w_i|^{p-1}, \end{aligned}$$

and thus by Hölder's inequality,

$$\begin{aligned} \|\vec{v} + \vec{w}\|_p^p &= \sum_{i=1}^{n} |v_i + w_i|^p \\ &\leq \sum_{i=1}^{n} |v_i|\,|v_i + w_i|^{p-1} + \sum_{i=1}^{n} |w_i|\,|v_i + w_i|^{p-1} \\ &\leq \|\vec{v}\|_p \left( \sum_{i=1}^{n} |v_i + w_i|^{pq-q} \right)^{\frac{1}{q}} + \|\vec{w}\|_p \left( \sum_{i=1}^{n} |v_i + w_i|^{pq-q} \right)^{\frac{1}{q}} \\ &= (\|\vec{v}\|_p + \|\vec{w}\|_p) \left( \sum_{i=1}^{n} |v_i + w_i|^p \right)^{\frac{1}{q}} = (\|\vec{v}\|_p + \|\vec{w}\|_p) \|\vec{v} + \vec{w}\|_p^{\frac{p}{q}}, \end{aligned}$$

and thus

$$\|\vec{v} + \vec{w}\|_p^{p(1-\frac{1}{q})} \leq \|\vec{v}\|_p + \|\vec{w}\|_p,$$

which is the desired result since $p(1 - \frac{1}{q}) = p \cdot \frac{1}{p} = 1$.

(c) Let $\vec{v} = (v_1, \cdots, v_n)^T \in \mathbb{R}^n$ and let $c \in \mathbb{R}$. Then

$$\begin{aligned} \|c\vec{v}\|_p &= \left( \sum_{i=1}^{n} |cv_i|^p \right)^{\frac{1}{p}} \\ &= \left( |c|^p \sum_{i=1}^{n} |v_i|^p \right)^{\frac{1}{p}} \\ &= |c|\|\vec{v}\|_p. \end{aligned}$$

Hence $\|\cdot\|_p$ is a vector norm.

15. As the vector $p$-norms are vector norms, the fact that entrywise norms are matrix norms directly follows from the linearity of the stacking operator:

(a) Let $A \in M_{m\times n}(\mathbb{R})$. Then

$$\|A\|_{p,e} = \|\mathcal{S}A\|_p \geq 0.$$

Let $\mathbf{0} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \in M_{m\times n}(\mathbb{R})$. Then

$$\|\mathbf{0}\|_{p,e} = \|\mathcal{S}\mathbf{0}\|_p = \|\vec{0}\|_p = 0,$$

where $\vec{0} = (\underbrace{0, \cdots, 0}_{mn \text{ zeros}})^T$ and $\|\cdot\|_p$ is the vector $p$-norm on $\mathbb{R}^{mn}$.

For any $A \in M_{m\times n}(\mathbb{R})$ such that $\|A\|_{p,e} = 0$,

$$\|\mathcal{S}A\|_p = 0 \text{ and thus } \mathcal{S}A = \vec{0},$$

which implies $A = \mathbf{0}$.

(b) For any $A, B \in M_{m\times n}(\mathbb{R})$,

$$\begin{aligned}\|A+B\|_{p,e} &= \|\mathcal{S}(A+B)\|_p \\ &= \|\mathcal{S}A + \mathcal{S}B\|_p \\ &\le \|\mathcal{S}A\|_p + \|\mathcal{S}B\|_p \\ &= \|A\|_{p,e} + \|B\|_{p,e}.\end{aligned}$$

(c) For any $A \in M_{m\times n}(\mathbb{R})$ and $c \in \mathbb{R}$,

$$\begin{aligned}\|cA\|_{p,e} &= \|\mathcal{S}(cA)\|_p \\ &= \|c\mathcal{S}A\|_p \\ &= |c|\,\|\mathcal{S}A\|_p \\ &= |c|\,\|A\|_{p,e}.\end{aligned}$$

On the other hand, we go to first principles for the induced norms.
Let $\|\cdot\|' : \mathbb{R}^n \to \mathbb{R}$ and $\|\cdot\|'' : \mathbb{R}^m \to \mathbb{R}$ be vector norms.
Define $\|\cdot\| : M_{m\times n}(\mathbb{R}) \to \mathbb{R}$ by

$$\|A\| = \sup_{\vec{x}\in\mathbb{R}^n, \vec{x}\ne\vec{0}'} \frac{\|A\vec{x}\|''}{\|\vec{x}\|'},$$

where $\vec{0}' = (\underbrace{0, \cdots, 0}_{n \text{ zeros}})^T$.

(a) Let $A \in M_{m\times n}(\mathbb{R})$. Since $\|\cdot\|'$ and $\|\cdot\|''$ are vector norms, for any $\vec{x} \in \mathbb{R}^n \setminus \{\vec{0}'\}$, $\|\vec{x}\|' > 0$, $\|A\vec{x}\|'' \ge 0$ and thus

$$\frac{\|A\vec{x}\|''}{\|\vec{x}\|'} \ge 0.$$

Taking supremum over all such $\vec{x}$, $\|A\| \ge 0$.

Let $\mathbf{0} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \in M_{m\times n}(\mathbb{R})$.

Then $\mathbf{0}\vec{x} = \vec{0}''$ for any $\vec{x} \in \mathbb{R}^n$, where $\vec{0}'' = (\underbrace{0, \cdots, 0}_{m \text{ zeros}})^T$, and thus

$$\|\mathbf{0}\| = \sup_{\vec{x}\in\mathbb{R}^n, \vec{x}\ne\vec{0}'} \frac{\|\mathbf{0}\vec{x}\|''}{\|\vec{x}\|'} = \sup_{\vec{x}\in\mathbb{R}^n, \vec{x}\ne\vec{0}'} \frac{\|\vec{0}''\|''}{\|\vec{x}\|'} = 0.$$

For any $A \in M_{m\times n}(\mathbb{R})$ such that $\|A\| = 0$, $\sup_{\vec{x}\in\mathbb{R}^n, \vec{x}\ne\vec{0}'} \frac{\|A\vec{x}\|''}{\|\vec{x}\|'} = 0$, i.e.

$$\text{for any } \vec{x} \in \mathbb{R}^n \setminus \{\vec{0}'\},\ \|A\vec{x}\|'' \le 0\|\vec{x}\|' = 0.$$

Combined with the fact that $\|\cdot\|''$ is a vector norm, $\|A\vec{x}\|'' = 0$ for any $\vec{x} \in \mathbb{R}^n$, which implies $A = \mathbf{0}$.

(b) For any $A, B \in M_{m\times n}(\mathbb{R})$,

$$\begin{aligned}\|A+B\| &= \sup_{\vec{x}\in\mathbb{R}^n, \vec{x}\ne\vec{0}'} \frac{\|(A+B)\vec{x}\|''}{\|\vec{x}\|'} \\ &\le \sup_{\vec{x}\in\mathbb{R}^n, \vec{x}\ne\vec{0}'} \frac{\|A\vec{x}\|'' + \|B\vec{x}\|''}{\|\vec{x}\|'} \\ &\le \sup_{\vec{x}\in\mathbb{R}^n, \vec{x}\ne\vec{0}'} \frac{\|A\vec{x}\|''}{\|\vec{x}\|'} + \sup_{\vec{y}\in\mathbb{R}^n, \vec{y}\ne\vec{0}'} \frac{\|B\vec{y}\|''}{\|y\|'} \\ &= \|A\| + \|B\|.\end{aligned}$$

(c) For any $A \in M_{m\times n}(\mathbb{R})$ and $c \in \mathbb{R}$,

$$\begin{aligned}\|cA\| &= \sup_{\vec{x}\in\mathbb{R}^n, \vec{x}\neq\vec{0}'} \frac{\|cA\vec{x}\|''}{\|\vec{x}\|'}\\ &= \sup_{\vec{x}\in\mathbb{R}^n, \vec{x}\neq\vec{0}'} \frac{|c|\|A\vec{x}\|''}{\|\vec{x}\|'}\\ &= |c| \sup_{\vec{x}\in\mathbb{R}^n, \vec{x}\neq\vec{0}'} \frac{\|A\vec{x}\|''}{\|\vec{x}\|'}\\ &= |c|\|A\|.\end{aligned}$$

16. (a) $A - \alpha B = \begin{pmatrix} 8-\alpha & 9-\alpha & 2-\alpha \\ 9-\alpha & 6-\alpha & 5-\alpha \\ 1-\alpha & -\alpha & 9-\alpha \end{pmatrix}$.

i. As $\|A-\alpha B\|_F \geq 0$ for any $\alpha \in \mathbb{R}$, minimizing $\|A-\alpha B\|_F$ is equivalent to minimizing $\|A-\alpha B\|_F^2$.

$$\begin{aligned}\|A-\alpha B\|_F^2 &= \alpha^2 + (1-\alpha)^2 + (2-\alpha)^2 + (5-\alpha)^2 + (6-\alpha)^2 + (8-\alpha)^2 + 3(9-\alpha)^2\\ &= 9\alpha^2 - 98\alpha + 373 = 9(\alpha - \frac{49}{9})^2 + \frac{956}{9},\end{aligned}$$

which is minimized at $\alpha = \frac{49}{9}$.

ii. Let

$$\begin{aligned}f(\alpha) &= \|A-\alpha B\|_{1,e}\\ &= |\alpha| + |1-\alpha| + |2-\alpha| + |5-\alpha| + |6-\alpha| + |8-\alpha| + 3|9-\alpha|.\end{aligned}$$

Since $|\cdot|$ is continuous on $\mathbb{R}$ and differentiable on $\mathbb{R}\setminus\{0\}$, $f$ is continuous on $\mathbb{R}$ and differentiable on $\mathbb{R}\setminus\{0,1,2,5,6,8,9\}$.
By first derivative test,

| $\alpha$ | $(-\infty,0)$ | $(0,1)$ | $(1,2)$ | $(2,5)$ | $(5,6)$ | $(6,8)$ | $(8,9)$ | $(9,\infty)$ |
|---|---|---|---|---|---|---|---|---|
| $f'(\alpha)$ | $-9$ | $-7$ | $-5$ | $-3$ | $-1$ | $1$ | $3$ | $9$ |

$f(\alpha) = \|A-\alpha B\|_{1,e}$ attains its minimum at $\alpha = 6$.

(b) $C - \alpha D = \begin{pmatrix} 7-\alpha & 7-\alpha & -\alpha & 9-\alpha \\ 3-\alpha & -\alpha & 8-\alpha & -\alpha \end{pmatrix}$.

i. As $\|C-\alpha D\|_F \geq 0$ for any $\alpha \in \mathbb{R}$, minimizing $\|C-\alpha D\|_F$ is equivalent to minimizing $\|C-\alpha D\|_F^2$.

$$\begin{aligned}\|C-\alpha D\|_F^2 &= 3\alpha^2 + (3-\alpha)^2 + 2(7-\alpha)^2 + (8-\alpha)^2 + (9-\alpha)^2\\ &= 8\alpha^2 - 68\alpha + 242 = 8(\alpha - \frac{17}{4})^2 + \frac{195}{2},\end{aligned}$$

which is minimized at $\alpha = \frac{17}{4}$.

ii. Let

$$\begin{aligned}g(\alpha) &= \|C-\alpha D\|_{1,e}\\ &= 3|\alpha| + |3-\alpha| + 2|7-\alpha| + |8-\alpha| + |9-\alpha|.\end{aligned}$$

Since $|\cdot|$ is continuous on $\mathbb{R}$ and differentiable on $\mathbb{R}\setminus\{0\}$, $g$ is continuous on $\mathbb{R}$ and differentiable on $\mathbb{R}\setminus\{0,3,7,8,9\}$.
By first derivative test,

| $\alpha$ | $(-\infty,0)$ | $(0,3)$ | $(3,7)$ | $(7,8)$ | $(8,9)$ | $(9,\infty)$ |
|---|---|---|---|---|---|---|
| $g'(\alpha)$ | $-8$ | $-2$ | $0$ | $4$ | $6$ | $8$ |

$g(\alpha)$ is minimized when $\alpha \in [3,7]$.

17. As $\|\cdot\| \geq 0$ for any $\alpha \in \mathbb{R}$, minimizing $\|\cdot\|_F$ is equivalent to minimizing $\|\cdot\|_F^2$.

(a) Note that $A_1 = A + \alpha \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ for some $\alpha \in \mathbb{R}$, and thus

$$\|A_1 - B\|_F^2 = \left\| A + \alpha \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} - B \right\|_F^2 = \left\| \begin{pmatrix} \alpha + 5 & \alpha - 1 \\ \alpha - 5 & \alpha - 3 \end{pmatrix} \right\|_F^2$$
$$= 4\alpha^2 - 8\alpha + 60 = 4(\alpha - 1)^2 + 56.$$

Hence $\|A_1 - B\|_F$ is minimized by $A_1 = A + \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 8 & 4 \\ 2 & 6 \end{pmatrix}$, and its minimum value is $2\sqrt{7}$.

(b) Note that the mean pixel value of $A$ is $\frac{7+3+1+5}{4} = 4$.
Hence $A_2 = \begin{pmatrix} 4 & 4 \\ 4 & 4 \end{pmatrix} + \beta \begin{pmatrix} 7-4 & 3-4 \\ 1-4 & 5-4 \end{pmatrix}$ for some $\beta \in \mathbb{R}$, and thus

$$\|A_2 - B\|_F^2 = \left\| \begin{pmatrix} 3\beta + 2 & -\beta \\ -3\beta - 2 & \beta - 4 \end{pmatrix} \right\|_F^2$$
$$= 20\beta^2 + 16\beta + 24 = 20(\beta + \frac{2}{5})^2 + \frac{104}{5}.$$

Hence $\|A_2 - B\|_F$ is minimized by $A_2 = \begin{pmatrix} 4 & 4 \\ 4 & 4 \end{pmatrix} - \frac{2}{5} \begin{pmatrix} 3 & -1 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} \frac{14}{5} & \frac{22}{5} \\ \frac{26}{5} & \frac{18}{5} \end{pmatrix}$, and its minimum value is $\frac{2\sqrt{130}}{5}$.

(c) Note that $A_3 = \begin{pmatrix} 4 & 4 \\ 4 & 4 \end{pmatrix} + \alpha \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \beta \begin{pmatrix} 3 & -1 \\ -3 & 1 \end{pmatrix}$ for some $\alpha, \beta \in \mathbb{R}$, and thus

$$\|A_3 - B\|_F^2 = \left\| \begin{pmatrix} 4 & 4 \\ 4 & 4 \end{pmatrix} + \alpha \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \beta \begin{pmatrix} 3 & -1 \\ -3 & 1 \end{pmatrix} - \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix} \right\|_F^2$$
$$= \left\| \begin{pmatrix} 2 + \alpha + 3\beta & \alpha - \beta \\ -2 + \alpha - 3\beta & -4 + \alpha + \beta \end{pmatrix} \right\|_F^2$$

$$= (\alpha + 3\beta + 2)^2 + (\alpha - \beta)^2 + (\alpha - 3\beta - 2)^2 + (\alpha + \beta - 4)^2$$
$$= 4\alpha^2 + 20\beta^2 - 8\alpha - 8\beta + 24 =: f(\alpha, \beta).$$

If $f$ is minimum at $(\alpha_0, \beta_0)$, $\frac{\partial f}{\partial \alpha}(\alpha_0, \beta_0) = \frac{\partial f}{\partial \beta}(\alpha_0, \beta_0) = 0$.
Hence by setting $\begin{cases} \frac{\partial f}{\partial \alpha} = 8\alpha - 8 = 0 \\ \frac{\partial f}{\partial \beta} = 40\beta - 8 = 0 \end{cases}$, we get $(\alpha_0, \beta_0) = (1, \frac{1}{5})$.
One can check that $\text{Hess}(f) \equiv \begin{pmatrix} 8 & 0 \\ 0 & 40 \end{pmatrix}$ is positive definite, and thus $f$ attains minimum at $(1, \frac{1}{5})$.
Hence $\|A_3 - B\|_F$ is minimized by $A_3 = \begin{pmatrix} 4 + 1 + 3 \cdot \frac{1}{5} & 4 + 1 - \frac{1}{5} \\ 4 + 1 - 3 \cdot \frac{1}{5} & 4 + 1 + \frac{1}{5} \end{pmatrix} = \begin{pmatrix} \frac{28}{5} & \frac{24}{5} \\ \frac{22}{5} & \frac{26}{5} \end{pmatrix}$, and its minimum value is $\sqrt{f(1, \frac{1}{5})} = \frac{4\sqrt{30}}{5}$.