2010 Autumn semester Pattern Information Processing
Topic 2. Image compression by orthogonal transformation

Session 6. (2) Orthogonal and unitary transformations of matrices

Image compression using vector expressions of images and transformation of vectors into principal components by orthogonal matrices were explained in the previous session. In this session, it is explained what corresponds to the above transformation in case that an image is expressed by a matrix. The concepts of orthogonal and unitary transformations and basis images are also explained.

Kronecker product and transformation of a matrix

Orthogonal transformation of vectors using vector expression of images was explained in the previous session. Digital image is, however, a two-dimensional array of pixel values. Thus the matrix expression is more suitable to an image.

We treated the transformation of an image vector x into a vector z by matrix P in the previous session, i. e.

$$z = P\boldsymbol{x}. \qquad (1)$$

We assume here that the column vectors $\boldsymbol{x}$ and z consist of m^2 elements and the matrix P is $m^2 \times m^2$.

We convert the image vector $\boldsymbol{x}$ into an image matrix. We separate the column vector $\boldsymbol{x}$ into m column subvectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_j, \ldots, \boldsymbol{x}_m$ of m elements, i. e.:

$$\boldsymbol{x} = \begin{pmatrix} \boldsymbol{x}_1 \\ \hline \vdots \\ \hline \boldsymbol{x}_j \\ \hline \vdots \\ \hline \boldsymbol{x}_m \end{pmatrix} \qquad (2)$$

For example, a column vector of 3^2 elements is separated into 3 column subvectors of 3 elements. We get an $m \times m$ matrix X by relocating these subvectors as column vectors along the row direction, i. e. :

$$X = \begin{pmatrix} \boldsymbol{x}_1 & \cdots & \boldsymbol{x}_j & \cdots & \boldsymbol{x}_m \end{pmatrix}. \qquad (3)$$

The matrix Z is obtained from the vector z in the same manner. It is assumed that the $m^2 \times m^2$ matrix P is expressed by the following $m \times m$ matrices C and R, as follows:

$$C = \begin{pmatrix} c_{11} & \cdots & c_{1m} \\ \vdots & \ddots & \vdots \\ c_{m1} & \cdots & c_{mm} \end{pmatrix}, \; R = \begin{pmatrix} r_{11} & \cdots & r_{1m} \\ \vdots & \ddots & \vdots \\ r_{m1} & \cdots & r_{mm} \end{pmatrix} \qquad (4)$$

$$P = \begin{pmatrix} r_{11}c_{11} & \cdots & r_{11}c_{1m} & & r_{1m}c_{11} & \cdots & r_{1m}c_{1m} \\ \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\ r_{11}c_{m1} & \cdots & r_{11}c_{mm} & & r_{1m}c_{m1} & \cdots & r_{1m}c_{mm} \\ & \vdots & & \ddots & & \vdots & \\ r_{m1}c_{11} & \cdots & r_{m1}c_{1m} & & r_{mm}c_{11} & \cdots & r_{mm}c_{1m} \\ \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\ r_{m1}c_{m1} & \cdots & r_{m1}c_{mm} & & r_{mm}c_{m1} & \cdots & r_{mm}c_{mm} \end{pmatrix} \qquad (5)$$

It means that the matrix P is composed by locating a copy of the matrix C at each element of the matrix R. This relationship is denoted as follows:

$$P = R \otimes C, \qquad (6)$$

and called that P is expressed by *Kronecker product* of R and C.

Using these expressions, the transformation of the vector $\boldsymbol{x}$ by the matrix P is expressed as follows:

$$Z = CXR', \qquad (7)$$

i. e. the transformation of the matrix X by the matrices C and R.

The above relationship is proved as follows: The vectors $\boldsymbol{x}, z$ contain the subvectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_j, \ldots, \boldsymbol{x}_m$ and $z_1, \ldots, z_j, \ldots, z_m$, respectively, as shown in Eq. (2). Let x_{ij}, z_{ij} be the ith elements of the jth subvectors $\boldsymbol{x}_j, z_j$, respectively. We get from Eq. (1) and (5) that

$$
\begin{aligned}
z_{ij} &= \left(r_{j1}c_{i1} \cdots r_{j1}c_{im} \; r_{jk}c_{i1} \cdots r_{jk}c_{im} \; r_{jm}c_{i1} \cdots r_{jm}c_{im} \right) \begin{pmatrix} x_{11} \\ \vdots \\ x_{m1} \\ \vdots \\ x_{1k} \\ \vdots \\ x_{mk} \\ \vdots \\ x_{1m} \\ \vdots \\ x_{mm} \end{pmatrix} \\
&= \sum_{k=1}^{m} r_{jk} \sum_{l=1}^{m} c_{il} x_{lk} \qquad (8)
\end{aligned}
$$

where k is an index of a subvector contained in $\boldsymbol{x}$, and l is an index of an element of the kth subvector $\boldsymbol{x}_k$. Since z_{ij} is an element in the ith row and jth column of the matrix Z, we get from Eq. (7) that

$$
\begin{aligned}
z_{ij} &= (CX)_{i\text{throw}} R'_{j\text{thcolumn}} \\
&= \left(\sum_{l=1}^{m} c_{il}x_{l1} \cdots \sum_{l=1}^{m} c_{il}x_{lk} \cdots \sum_{l=1}^{m} c_{il}x_{lm} \right) \begin{pmatrix} r_{j1} \\ \vdots \\ r_{jk} \\ \vdots \\ r_{jm} \end{pmatrix} \\
&= \sum_{k=1}^{m} r_{jk} \sum_{l=1}^{m} c_{il} x_{lk}. \qquad (9)
\end{aligned}
$$

Equations (8) and (9) show that Eq.(1) and Eq. (7) are equivalent.

The inner product of two columns of P is as follows:

$$
\begin{aligned}
& \sum_{i=1}^{m} r_{1k}c_{il} \cdot r_{1k'}c_{il'} + \cdots \\
& \quad + \sum_{i=1}^{m} r_{jk}c_{il} \cdot r_{jk'}c_{il'} + \cdots + \sum_{i=1}^{m} r_{mk}c_{il} \cdot r_{mk'}c_{il'} \\
&= \sum_{j=1}^{m} \sum_{i=1}^{m} r_{jk}c_{il} \cdot r_{jk'}c_{il'} \\
&= \sum_{j=1}^{m} r_{jk}r_{jk'} \sum_{i=1}^{m} c_{il}c_{il'} \qquad (10)
\end{aligned}
$$

Assuming that P is orthonormal, the inner product is 1 if $k = k'$ and $l = l'$, and 0 otherwise. Since the former summation in the bottom row of Eq. (10) means the product of the kth column and k'th column of the matrix R, and the latter summation means the product of the lth column and l'th column of the matrix C, the matrix P is orthonormal if both R and C are orthonormal.

We conclude from the above discussion that the transformation of the image vector $\boldsymbol{x}$ by the orthonormal matrix P is expressed by Eq. (7) in case of converting the vector $\boldsymbol{x}$ to the matrix X, if P is expressed by the Kronecker product of the orthonormal matri-

[1]The symbols "C" and "R" suggest the initials of Column and Row.

ces C and R. Equation (7) indicates[1] that the matrix C operates on the columns of the matrix X, and the matrix R operates on the rows of X. It follows that "P can be expressed by the Kronecker product of C and R" means "The operation of P can be separated into the operation on the columns and that on the rows of X." This is called that P is *separable*.

Orthogonal and unitary transformations of matrices

The objective of orthogonal transformation of an image is separating the image into "more important components" and "less important components," as explained in the previous session. It was explained in the previous session how to reduce the latter elements of the vector z in the transformation of Eq. (1). If we employ the expression of Eq. (7), explained in this session, it is known that there exist such C and R that Z becomes diagonal. The method to find such C and R is known as *singular value decomposition* (SVD). If the $m \times m$ matrix Z is diagonal, there are only m nonzero elements in Z. It indicates that the data compression similar to that in the previous session is achieved.

However, different C's and R's must be chosen for different X's by SVD. It indicates that the SVD cannot achieve the transmission of images with a small data amount by operating a *common* transformation on various images. This problem is the same in the case of KL transformation, explained in the previous session, since the covariance matrix of the image set under consideration must be known.

To avoid this problem, we consider such C and R that are independent on the input image X and that the transformed image Z contains as many elements that are zero or nearly zero as possible. Assuming that $C = R$ since it is very rare that different operations should be applied to rows and columns of an image, Eq. (7) is rewritten to

$$Z = RXR'. \tag{11}$$

Since R is orthonormal, $RR' = I$. Thus the inverse transformation is

$$X = R'ZR. \tag{12}$$

Such operation is called orthogonal transformation of an image. If we assume that the elements of R are complex, we select such R that $RR'^{*} = I$. The symbol * indicates the complex conjugate. Such matrix is called unitary matrix, and the transformation pair in Eq. (11) and (12) is called *unitary transformation*.

Basis Image

Consider the decomposition of Z into the summation of m^2 matrices each of which contains only one nonzero element, i. e.

$$Z = \begin{pmatrix} z_{11} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} + \begin{pmatrix} 0 & z_{12} & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} + \cdots + \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & z_{mm} \end{pmatrix}. \tag{13}$$

If we express R using row vectors as follows:

$$\boldsymbol{r}'_j = \left(r_{j1} \cdots r_{jk} \cdots r_{jm}\right),\ R = \begin{pmatrix} \boldsymbol{r}'_1 \\ \vdots \\ \boldsymbol{r}'_j \\ \vdots \\ \boldsymbol{r}'_m \end{pmatrix}, \tag{14}$$

Eq. (12) is rewritten to the summation of the matrices composed by outer products of the rows of R, i. e. the columns of R', as follows:

$$
\begin{aligned}
X &= R' \begin{pmatrix} z_{11} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} R + R' \begin{pmatrix} 0 & z_{12} & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} R + \cdots + R' \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & z_{mm} \end{pmatrix} R \\
&= R' z_{11} \begin{pmatrix} \boldsymbol{r}'_1 \\ \boldsymbol{0}' \\ \vdots \\ \boldsymbol{0}' \end{pmatrix} + R' z_{12} \begin{pmatrix} \boldsymbol{r}'_2 \\ \boldsymbol{0}' \\ \vdots \\ \boldsymbol{0}' \end{pmatrix} + \cdots + R' z_{12} \begin{pmatrix} \boldsymbol{0}' \\ \boldsymbol{0}' \\ \vdots \\ \boldsymbol{r}'_m \end{pmatrix} \\
&= z_{11} \boldsymbol{r}_1 \boldsymbol{r}'_1 + z_{12} \boldsymbol{r}_1 \boldsymbol{r}'_2 + \cdots + z_{mm} \boldsymbol{r}_m \boldsymbol{r}'_m \\
&= \sum_{i=1}^{m} \sum_{j=1}^{m} z_{ij} \boldsymbol{r}_i \boldsymbol{r}'_j
\end{aligned} \tag{15}
$$

These outer products are called *basis images.* In this expression, we can consider the elements z_{ij} of the transformed image to be the coefficients of the basis images, i. e. we can consider the original image X to be the summation of every basis image multiplied by z_{ij}. If we assume the transmitter and the receiver commonly know the basis images in advance, the image compression is achieved by transmitting the nonzero z_{ij} ' s only, if only a certain number of z_{ij} are regarded as almost zero. The image X is expressed almost exactly by the basis images corresponding to the nonzero z_{ij} ' s.

It is noticed that the above discussion is equivalent to the discrete Fourier transformation explained in the topic 1 if we replace the basis image with the exponential function. The two-dimensional Fourier transformation is one of unitary transformations, and in this case combinations of sinusoidal waves of various frequencies appear as basis images. If the image does not consist of high frequency components, the basis images for high frequency waves are not used, i. e. z_{ij} ' s corresponding to those basis images do not have to be transmitted.

Various unitary transformations for the image compression have been proposed. The next session, the final of this topic, the discrete Fourier transformation as a unitary transformation, and the discrete cosine transformation for the JPEG image compression, will be explained.