线性代数

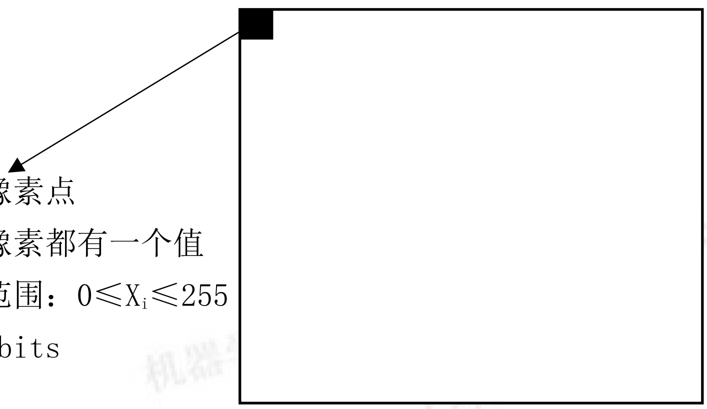
-32 课 基变换和图像压缩

一、知识概要

本节主题是线性变换与矩阵的关联，从图像压缩与信号处理的应用引入，介绍几种方便的基向量：傅里叶，小波。最后从代数角度大体上介绍了基变换与变换矩阵的关系。

二．图像处理

首先我们假设有一个 512*512 的黑白图像，那么它的大体性质如下：



一个像素点
每个像素都有一个值
变化范围：$0 \leqslant X_i \leqslant 255$
占 8 bits

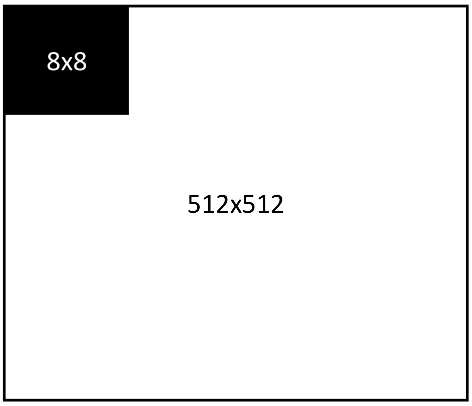
由此看来，每个图像都可以使用一个向量 x 来表示，其中的各个分量代表图像中的每一个像素点。此时，$x \in R^n$，n = 512*512。这既是我们得到的信息，将图像抽象为一个长度为 n 的向量。如果是彩色图像长度就是三倍，就是 3*512*512。因为我们需要三个值来代表颜色。

我们使用 JPEG 方式压缩图像。就是利用基变换来压缩储存空间。怎么做呢？具体来说，一个图像中总有颜色相近的点，或者类似的点，而我们在压缩之前是使用标准基储存图像，其实我们没必要这么做，比如蓝衬衫，黑板这类颜色在很大一部分上是相近的。那么我们只要使用$\begin{vmatrix}1\\1\\\cdot\\\cdot\\\cdot\\1\end{vmatrix}$来储存这部分信息即可。不需要再用标准基来细致入微的储存相近的信息。由此我们看出，怎样选择储存信息的基，是很重要的一个问题。

三．两个重要的基

2.1 傅里叶基

JPEG 处理图像压缩的方法就是先将图像分块，在使用傅里叶基进行处理，最后进行压缩。

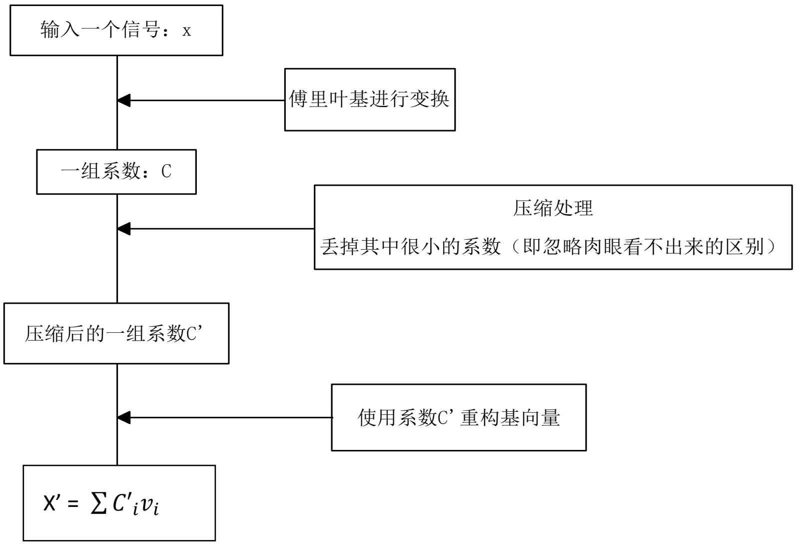


如上图，首先将图片划分为若干个 8*8 区域，每个区域中有 64 个元素，再使用傅里叶基进行变换。

8*8 的图像傅里叶基如下：

$$\begin{bmatrix}1\\1\\1\\1\\1\\1\\1\\1\end{bmatrix}\begin{bmatrix}1\\w\\w^2\\w^3\\w^4\\w^5\\w^6\\w^7\end{bmatrix}\cdots\cdots\begin{bmatrix}1\\w^7\\w^{14}\\w^{21}\\w^{28}\\w^{35}\\w^{42}\\w^{49}\end{bmatrix}$$

整个处理流程如下：



如上，整个处理过程中，第一步傅里叶基的变换过程是无损处理，而第二部的压缩过程是有损处理，最后导致 C’中很多项都是 0，需要储存的仅剩下很少的几项，这个过程中我们完成了压缩。

2.2 小波

小波也是一组很好的基，在 8*8 的情况下，其基为：

$$\begin{bmatrix}1\\1\\1\\1\\1\\1\\1\\1\end{bmatrix}\begin{bmatrix}1\\1\\1\\1\\-1\\-1\\-1\\-1\end{bmatrix}\begin{bmatrix}1\\1\\-1\\-1\\0\\0\\0\\0\end{bmatrix}\begin{bmatrix}0\\0\\0\\0\\1\\1\\-1\\-1\end{bmatrix}\begin{bmatrix}1\\-1\\0\\0\\0\\0\\0\\0\end{bmatrix}\begin{bmatrix}0\\0\\1\\-1\\0\\0\\0\\0\end{bmatrix}\begin{bmatrix}0\\0\\0\\0\\1\\-1\\0\\0\end{bmatrix}\begin{bmatrix}0\\0\\0\\0\\0\\0\\1\\-1\end{bmatrix}$$

假设上面小波基向量构成一个矩阵 W，我们这时就可以将向量 p 变化为：
$p = c_1w_1 + c_2w_2 + c_3w_3 + \cdots\cdots c_8w_8$，或者写为：p = WC，这样一来，即可得到：$C = W^{-1}P$

2.3 总结

以上两种基之所以可以成为比较方便的基，原因如下：

（1） 逆矩阵可以被快速求出，之前我们学习了快速傅里叶变换 FFT，而还存在着一种快速小波变换，因为我们注意到，小波的各个向量之间互相正交，那么我们就可以将它们化为标准正交基，联系之前的知识，标准正交基的逆就是其转置，故而其逆矩阵可以快速求出。

（2） 具有良好的压缩性，就以小波为例，其中的后面部分$\begin{bmatrix}1\\-1\\0\\0\\0\\0\\0\\0\end{bmatrix}\begin{bmatrix}0\\0\\1\\-1\\0\\0\\0\\0\end{bmatrix}\begin{bmatrix}0\\0\\0\\0\\1\\-1\\0\\0\end{bmatrix}\begin{bmatrix}0\\0\\0\\0\\0\\0\\1\\-1\end{bmatrix}$，如果我们在压缩过程中丢掉这些向量，那么我们损失的只是很少的点，而如果我们的基中元素太多，在我们压缩时，丢掉的点所代表的就会是很多数据，这样会使压缩品质不够良好。所以我们认为，少量的基向量就能接近信号，重现图像的基才是性质良好的。

四．基变换

4.1 坐标角度的基变换

从上面的应用中我们可以看出来，基变换的意义十分重要。这里进行一下复习，我们令 W 的列向量是新的基向量。x 是旧基下的一个向量，而 c 是新基下的对应得到的向量，则：

$x = Wc$　:矩阵 W 给出了一种基变换形式(即是上节中的矩阵 A)。

上面关于小波的介绍中就使用了这种方法确定向量在新旧基下坐标转换的关系。

4.2 线性变换矩阵角度的基变换

接下来我们通过线性变换的角度再谈一谈，若一线性变换 T，是从 8 维到 8 维的：$R^8 \to R^8$，那么：

一组基 v_1，v_2，v_3……v_8，线性变换矩阵为 A

另外一组基 w_1，w_2，w_3……w_8，线性变换矩阵为 B

首先注意这里的背景，这里是在讨论两种情况，即线性变换输入输出空间相同，也就是 A 作用前输入向量的基为 v_i，则 A 作用后输出空间基还是 v_i。B 也是一样。线性变换方式没有变化，只是在不同基下对应的坐标不同，线性变换矩阵不同。我们这里举这个例子是为了从线性变换矩阵的角度认识：不同基下，同一个线性变换对应的线性变换矩阵会有什么关系。

直接给出结论：如果是同一种线性变换，在不同基下对应的线性变换矩阵为 A，B，则 A 与 B 相似。即：$B = M^{-1}AM$　（其中的 M 可以与 W 相同）

4.3 总结

对于一个向量来说，变换一组基，发生两件事：

- 第一，不同的基对应了不同的坐标，$x = Wc$ 即是新旧坐标之间的关系。(4.1)
- 第二，每一中情况下的线性变换矩阵也变了，不同的基下得到了新矩阵，这里得到的 A，B 之间有相似关系：

$$B = M^{-1}AM \qquad \text{（其中的 M 可以与 W 相同）} \qquad (4.2)$$

以 A 为例，我们复习一下线性变换过程：

$$x = c_1v_1 + c_2v_2 + c_3v_3 + \cdots\cdots + c_8v_8$$

则变换 T 得到：

$T(x) = c_1T(v_1) + c_2T(v_2) + c_3T(v_3) + \cdots\cdots + c_8T(v_8)$

这之后我们想要的就是 T 作用在这些基上的结果：

$T(v_1) = a_{11}v_1 + a_{21}v_2 + a_{31}v_3 + \cdots\cdots + a_{81}v_8$

$T(v_2) = a_{12}v_1 + a_{22}v_2 + a_{32}v_3 + \cdots\cdots + a_{82}v_8$

……………

$T(v_8) = a_{18}v_1 + a_{28}v_2 + a_{38}v_3 + \cdots\cdots + a_{88}v_8$

而 A 就是上述 a 构成的矩阵：

$$[A] = \begin{bmatrix} a_{11} & a_{21}\cdots & a_{81} \\ a_{12} & a_{22}\cdots & a_{82} \\ \vdots & \vdots\ \vdots & \vdots \\ a_{18} & a_{28}\cdots & a_{88} \end{bmatrix}$$

特别的，如果 v_1，v_2，v_3……v_8 都是 x 的特征向量，那么我们这一变换得到的矩阵 A 就是对角矩阵，矩阵对角线上排列特征值。

五. 学习感悟

本节内容主要通过线性变换与基变换介绍了其应用：图像压缩，这部分的介绍都比较概括，教授的主要目的是让我们了解这些东西都可以用来做什么以及它们都具有怎样的性质。主要注意的就是线性变换与矩阵之间的关系，以及不同基下的变换会有什么特点。