ESTR1005：线性代数 笔记

> $Linear\ Algebra\ for\ Engineers$

1. 线性方程

> $Linear\ Equations\ in\ Linear\ Algebra$

1.1 线性方程组

> $Systems\ of\ Linear\ Equations$

① 线性方程

> $Linear\ Equation$

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

> b and the coefficients a_1 , ··· , a_n are real or **complex** numbers；
> 某些 **coefficients 可以为 0**，不用从 x_1 开始

$4x_1 - 5x_2 = x_1x_2$

$x_2 = 2\sqrt{x_1} - 6$

不是线性方程

② 线性方程组

> $A\ system\ of\ linear\ equations\ /\ linear\ system$

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned}$$

> 类似地，某些 **coefficients 可以为 0**，不用从 x_1 开始

③ 解

> $system$ 的 $solution$: 使 $system$ 中每个 $equation$ 都为 $true\ statement$ 的一组 $numbers$

这里实际上延申了 $solution$ 的概念。从学过的 $equation$ 的解，延申到 $system$ 的解

一个 $system$ 的解有三种情况：

(1) $no\ solution$；(2) $exactly\ one\ solution$；(3) $infinitely\ many\ solutions$.

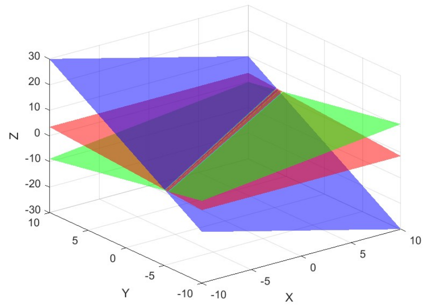
解释一下为什么不可能有两解 / 三解 等情况：

$system$ 本质上是 n 维空间中，若干 $n-1$ 维物体的联立。

例如下面这个 $system$

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned}$$

是三维空间中，三个二维物体（平面）的联立。这三个平面的公共部分就是该 $system$ 的 $solution\ set$



假设已找到某 $system$ 的两个解。

首先，这两个解不可能是若干零维物体（点）的公共部分；

对于 $\geqslant$ 一维的物体（直线、平面、空间$\cdots$），两点确定一条直线，有两解就有无数解

相容

> $consistent$

若线性方程组 $A\mathbf{x} = \mathbf{b}$ 有解（一解或无穷解），则称 $A\mathbf{x} = \mathbf{b}$ 是相容的（$consistent$）；

若线性方程组 $A\mathbf{x} = \mathbf{b}$ 无解，则称 $A\mathbf{x} = \mathbf{b}$ 是不相容的（$inconsistent$）.

④ 解集

> $solution\ set$: $all\ possible\ solutions$ 的集合

同样也是延申到 $system$ 的概念

⑤ 方程组等价

> $equivalent$: Two linear systems are called **equivalent** if they have the same solution set.
> (or each of them is the subset of the other one, 相互包含)

⑥ 矩阵

> $matrix\ notation$ (复数：$matrices$)

一个 $system$ 的主要信息可以用 $matrix$ 来表示 . 如下列 $system$

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned}$$

有两类矩阵：

系数矩阵

$$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ -4 & 5 & 9 \end{bmatrix}$$

> $coefficient\ matrix\ /\ matrix\ of\ coefficients$
> 不含等号右边的常数

增广矩阵

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{bmatrix}$$

> $augmented\ matrix$
> 含等号右边的常数

维数

> the **size** of the matrix : $m \times n$ matrix m 行 n 列矩阵

注意与后面学的 “维度(dimension)” 区分。

⑦ 初等行变换

> $elementary\ row\ operations$

倍加变换（replacement）	把某一行换成它本身与另一行倍数的和.
对换变换（interchange）	把两行对换.
倍乘对换（scaling）	把某一行的所有元素乘以同一个非零数.

> 行变换是可逆的（$reversible$）

⑧ 行等价

> 行变换前后，矩阵是行等价的（$row\ equivalent$）

若两个 $linear\ system$ 的 $augmented\ matrices$ 是行等价的，则它们有相同的 $solution\ set$.

1.2 行化简与阶梯形矩阵

> $Row\ Reduction\ and\ Echelon\ Forms$

① 定义

项 / 元（entry）	矩阵中的元素.
非零行 / 非零列（nonzero row / nonzero column）	至少包含一个非零元素的行 / 列.
先导元素（leading entry）	非零行最左边的非零元素.
阶梯形矩阵（echelon form）	零行沉底；先导元素逐行靠右（可以不相邻）.
简化阶梯形矩阵（reduced echelon form）	先导元素为 1 且上下均为 0 的阶梯形.

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

阶梯形，但非简化

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 3 & 4 \\ 0 & 0 & 0 & 4 \end{bmatrix}$$

阶梯形（可以不相邻），但未简化

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 2 & 3 \end{bmatrix}$$

非阶梯形（第三行先导元素未靠右）

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & 2 & 3 \end{bmatrix}$$

非阶梯形（零行未沉底）

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

阶梯形，但未简化（第二列和第三列，先导元素都不是所在列唯一的非零元素）

$$\begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 4 \end{bmatrix}$$

简化阶梯形

$$\begin{bmatrix} 1 & 0 & 0 & 2 & 5 \\ 0 & 1 & 0 & 3 & 6 \\ 0 & 0 & 1 & 4 & 7 \end{bmatrix}$$

简化阶梯形

$$\begin{bmatrix} 1 & 0 & 0 & \frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{3} \\ 0 & 0 & 1 & \frac{1}{4} \end{bmatrix}$$

简化阶梯形（非先导元素可以为分数）

② Theorem 1.2.1

简化阶梯形矩阵的唯一性

> $Uniqueness\ of\ the\ Reduced\ Echelon\ Form$

任意非零矩阵可能通过 $row\ operations$ 转换成不同的 $echelon\ matrix$，但是 $reduced\ echelon\ form$ 只有一个

证　利用 4.3 节的思想，即行等价矩阵的列具有完全一样的线性相关关系.

行化简算法说明至少存在一个这样的矩阵 U. 假设 A 行等价于简化阶梯形矩阵 U 和 V，U 的行中最左边非零元素是“主元”1，称这类主元 1 的位置是一个主元位置，并且包含它的列称为主元列（这个定义仅用于阶梯特征的 U 和 V，并且假设简化阶梯形不唯一）.

U 和 V 的主元列恰恰是与它们左边的列线性无关的非零列（这个条件自动满足第一列是非零的）. 由于 U 和 V 行等价（两个矩阵都行等价于 A），故它们的列有相同的线性相关关系，因此，U 和 V 的主元列出现在同样位置. 如果有 r 个这样的列，则因为 U 和 V 是简化阶梯形，所以它们的主元列是 $m \times n$ 单位矩阵的前 r 列，这样对应的 U 和 V 的主元列是相等的.

最后，考虑 U 的任意非主元列，例如列 j，这个列或者是零或者是左边主元列的线性组合（因为这些主元列是第 j 列左边的列生成的空间的一个基）. 两种情形下，对第 j 个元素为 1 的 $\mathbf{x}$ 都可写成 $U\mathbf{x} = \mathbf{0}$. 那么 $V\mathbf{x} = \mathbf{0}$，这说明 V 的第 j 列或者是零或者是与它左边 V 的主元列相同的线性组合. 由于 U 和 V 对应的主元列是相等的，因此 U 和 V 的第 j 列也相等，这个结果对 U 和 V 所有非主元列也成立，因而 $V=U$，这就证明了 U 是唯一的. ■

> 证明很恶心，可以不看

③ 主元位置和主元列

主元位置

> $pivot\ position$，矩阵对应 $reduced\ echelon\ form$ 中先导元素 1 的位置。
>
> 注意是矩阵本身的位置，不是简化阶梯形的位置，简化阶梯形只是帮助定位。

主元列

> $pivot\ column$，$pivot\ position$ 所在列。
>
> 同样也是矩阵本身的列，可能并不含有先导元素。

④ 行化简算法

> $row\ reduction\ algorithm$

提前说明：此方法可解线性方程组 / 矩阵方程，后面会具体介绍。

下面先以一个实例来说明它的具体步骤。

实例演示：用初等行变换化简下列矩阵。

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix}$$

向前步骤：化简为 $echelon\ form.$

> $forward\ phase$

$Step\ 1$：选中最左非零列，$interchange$ 使 $top\ entry\,(\,pivot\ position\,)$ 非零。

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix} \Rightarrow \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

$Step\ 2$：$replacement$ 使 $pivot$ 下全为零。

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix} \Rightarrow \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

$Step\ 3$：遮住 $pivot$ 所在行及其上所有行，重新进行 $Step\ 1$ 和 $Step\ 2$，直到所有非零行修改完毕。

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix} \Rightarrow \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

向后步骤：进一步化简为 $reduced\ echelon\ form.$

> $backward\ phase$

$Step\ 4$：从右往左，$scaling$ 使 $pivot$ 为 1，再 $replacement$ 使 $pivot$ 上全为零。

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 2 & -4 & 4 & 0 & -14 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 3 & 0 & -6 & 9 & 0 & -72 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

C语言代码：

```
#include <stdio.h>
#define acceptable_error 0.001

int max = 10, m, n, pivot_row = 0, pivot_column = 0;

float absolute_value(float a, float b) {
  float abs = 0;

  if (a >= b)
    abs = a - b;
  else
    abs = b - a;

  return abs;
}

void print_matrix(double **matrix, int m, int n, int max) {
  for (int i = 0; i < m; i++) {
    for (int j = 0; j < n; j++)
      printf("%6.2f ", *((double *)matrix + max * i + j));
    printf("\n");
  }
  return;
}

void step_1(double **matrix, int m, int n, int max) {
  int flag = 0;
  for (int j = 0; j < n; j++) {
    for (int i = pivot_row; i < m; i++)
      if (*((double *)matrix + max * i + j) != 0) {
        pivot_column = j;
        if (i != pivot_row) {
          for (int k = pivot_column; k < n; k++) {
            double temp = *((double *)matrix + max * i + k);
            *((double *)matrix + max * i + k) =
                *((double *)matrix + max * pivot_row + k);
            *((double *)matrix + max * pivot_row + k) = temp;
          }
        }
        flag++;
        break;
      }
    if (flag != 0)
      break;
  }
  if (flag == 0) {
    pivot_row = -1;
    pivot_column = -1;
  }
  return;
}

void step_2(double **matrix, int m, int n, int max) {
```

```
  if (pivot_column == -1) {
    printf("阶梯形为：\n");
    return;
  } else {
    for (int i = pivot_row + 1; i < m; i++)
      if (*((double *)matrix + max * i + pivot_column) != 0) {
        double rate = *((double *)matrix + max * i + pivot_column) /
                      *((double *)matrix + max * pivot_row + pivot_column);
        for (int j = pivot_column; j < n; j++)
          *((double *)matrix + max * i + j) -=
              rate * (*((double *)matrix + max * pivot_row + j));
      }
    pivot_row++;
    return;
  }
}

void step_3(double **matrix, int m, int n, int max) {

  for (int i = m - 1; i >= 0; i--)
    for (int j = 0; j < n; j++)
      if (absolute_value(*((double *)matrix + max * i + j), 0) >
          acceptable_error) {
        float divide = *((double *)matrix + max * i + j);
        for (int k = 0; k < n; k++) {
          *((double *)matrix + max * i + k) =
              *((double *)matrix + max * i + k) / divide;
        }

        for (int k = i - 1; k >= 0; k--) {

          double rate = *((double *)matrix + max * k + j) /
                        *((double *)matrix + max * i + j);

          for (int t = j; t < n; t++)
            *((double *)matrix + max * k + t) -=
                rate * (*((double *)matrix + max * i + t));
        }

        break;
      }
  return;
}

int main(void) {
  double matrix[max][max];
  printf("请输入 m x n 矩阵对应参数：(m,n为小于等于%d的正整数)\n", max);
  printf("m=");
  scanf("%d", &m);
  printf("n=");
  scanf("%d", &n);
  printf("\n请输入需要化简的 %d x %d 矩阵：\n", m, n);
  for (int i = 0; i < m; i++)
    for (int j = 0; j < n; j++)
      scanf("%lf", &matrix[i][j]);
  printf("\n化简矩阵：\n");
  print_matrix((double **)matrix, m, n, max);

  while (pivot_row != -1) {
    step_1((double **)matrix, m, n, max);
    printf("\n");
    print_matrix((double **)matrix, m, n, max);
    printf("\n");
    step_2((double **)matrix, m, n, max);
    printf("\n");
    print_matrix((double **)matrix, m, n, max);
  }
  printf("\n");
  step_3((double **)matrix, m, n, max);
  printf("简化阶梯形为：\n\n");
  print_matrix((double **)matrix, m, n, max);

  return 0;
}
```

⑤ 基本变量和自由变量

> $basic\ variables\ \&\ free\ variables$

将行化简算法应用于增广矩阵，得到线性方程组解集的一种显式表示（$an\ explicit\ description$）。

例如，某个 $augmented\ matrix$ 被化简为如下 $reduced\ echelon\ form$：

$$\begin{bmatrix} 1 & 0 & -5 & 1 \\ 0 & 1 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

对应方程组为

$$\begin{aligned} x_1 - 5x_3 &= 1 \\ x_2 + x_3 &= 4 \\ 0 &= 0 \end{aligned}$$

对应矩阵 $pivot\ columns$ 的变量 x_1，x_2 称为 基本变量，其他（x_3）称为 自由变量。

说明：线性方程组解到这就算解出来了（x_1，x_2 与 x_3 的关系很明显了）。

但实际解题中，解集习惯用另一种写法：

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5x_3 + 1 \\ -x_3 + 4 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} 5 \\ -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 4 \\ 0 \end{bmatrix}$$

> x_3 任取，因此第二项有无数种可能。后面"矩阵方程"会详细讨论。

⑥ 解集的参数表示

> $parametric\ descriptions\ of\ solutions\ sets$

解集的参数表示：自由变量作参数，表示基本变量。

方程组相容且有自由变量时，解集有多种参数表示。

方程组不相容时，解集是空集，无论是否有自由变量，解集均无参数表示。

解方程组：求出解集的参数表示，或确定它无解。

⑦ Theorem 1.2.2

存在与唯一性定理

> $Existence\ and\ Uniqueness\ Theorem$

线性方程组相容 $\Leftrightarrow$ 增广矩阵最右列不是主元列 $\Leftrightarrow$ 增广矩阵的阶梯形（注意，不需要简化）没有形如 $[\,0 \quad \cdots \quad 0 \quad b\,]$, $b \neq 0$ 的行

线性方程组相容时，解集有 2 种可能：

（1）没有自由变量时，有唯一解；

（2）有自由变量时，有无穷多解。

⑧ 用行化简算法解线性方程组

> $Using\ Row\ Reduction\ to\ Solve\ a\ Linear\ System$

$Step\ 1$：写出方程组的增广矩阵。

$Step\ 2$：行化简为阶梯形，确定是否相容：不相容则停止（解集为空集），相容则继续。

$Step\ 3$：行化简为简化阶梯形。

$Step\ 4$：根据简化阶梯形重新写出方程组。

$Step\ 5$：把 $Step\ 4$ 中每个非零方程改写为用自由变量表示基本变量（主元列对应的变量）的形式。

1.3 向量方程

> $vector\ equations$

① 向量定义

> $vector\ /\ column\ vector$

只有一列的矩阵叫做列向量，习惯直接叫做向量。

> 注意：线代中研究的向量和矩阵并没有本质区分，也不太关注它的"方向"，定义会有点奇怪。
>
> 后续可能会用到一些几何解释，但在线代中几何仅作为辅助，方便理解，不应把它当作重点，
>
> 因为高维的几何空间难以想象，所以重点在于对线代各种定理和"线性变换"的理解。

$W = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$ is a vector in $\mathbb{R}^2$, where w_1 and w_2 are real numbers.

② 向量运算

> $equal\ /\ sum\ /\ scalar\ multiple$

按常理理解即可。

两个向量 $equal \Leftrightarrow$ 对应项相同

两个向量的 $sum \Leftrightarrow$ 对应项相加

> 隐含条件：维度相同。维度不同无法比较/相加。

标量乘法/数乘（$scalar\ multiple$）：给定实数 c 和向量 $\mathbf{u}$，把 $\mathbf{u}$ 的每一项乘以 c，所得向量记为 $c\mathbf{u}$.

$zero\ vector$：各项均为 0 .

③ 几何解释

$\mathbb{R}^2$：平面中所有点的集合

$\mathbf{u}+\mathbf{v}$：平行四边形法则

$\mathbb{R}^n$：有序 n 元组的集合 ($ordered\ n-tuples$)

④ 线性组合

$linear\ combination$ 线性组合

Given vectors $\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_p} \in \mathbb{R}^n$ and scalars $c_1, c_2, \cdots, c_p$.

$\mathbf{y} = c_1\mathbf{v_1} + c_2\mathbf{v_2} + \cdots + c_p\mathbf{v_p}$

is called a linear combination of $\mathbf{v_1}, \cdots, \mathbf{v_p}$ with weights $c_1, \cdots, c_p$.

> 隐藏条件：the vectors are in the same dimension

$weights$ 权重

> weights 可以是任意实数，包括 0

⑤ 向量方程

向量方程 $x_1\mathbf{a_1} + x_2\mathbf{a_2} + \cdots + x_n\mathbf{a_n} = \mathbf{b}$

和增广矩阵为 $[\ \mathbf{a_1} \quad \mathbf{a_2} \quad \cdots \quad \mathbf{a_n} \quad \mathbf{b}\ ]$ 的线性方程组有相同的解集。

> 特别地，$\mathbf{b}$ 可表示为 $\mathbf{a_1}, \mathbf{a_2}, \cdots, \mathbf{a_n}$ 的线性组合，当且仅当该线性方程组有解。

⑥ 张成

> $span$

If $\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_p}$ are in $\mathbb{R}^n$, then the set of all linear combinations of $\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_p}$ is denoted by $Span\{\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_p}\}$ and is called **the subset of** $\mathbb{R}^n$ spanned (or generated) by $\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_p}$.

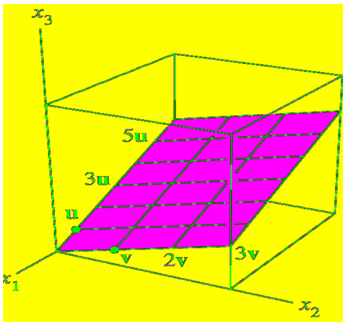
> That is, $Span\{\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_p}\}$ is the collection of all vectors that can be written in the form $c_1\mathbf{v_1} + c_2\mathbf{v_2} + \cdots + c_p\mathbf{v_p}$ with $c_1, c_2, \cdots, c_p$ scalars.

注意，$\mathbb{R}^n$ 中向量的张成是 $\mathbb{R}^n$ 的子集，暗示张成的维度不可能超过 n.

> 举个例子：平面中取一些向量，不可能生成三维空间。

$\mathbb{R}^3$ 中，给定非零向量 $\mathbf{v}$，$Span\{\mathbf{v}\}$ 构成 $\mathbf{v}$ 所在直线

给定非零向量 $\mathbf{u}, \mathbf{v}$（$\mathbf{u}$，$\mathbf{v}$ 不共线），$Span\{\mathbf{u}，\mathbf{v}\}$ 构成 $\mathbf{u}, \mathbf{v}$ 所在平面



1.4 矩阵方程 $A\mathbf{x} = \mathbf{b}$

> THE MATRIX EQUATION $A\mathbf{x} = \mathbf{b}$

① $A\mathbf{x}$ 定义

$$A\mathbf{x} = \begin{bmatrix} \mathbf{a_1} & \mathbf{a_2} & \cdots & \mathbf{a_n} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a_1} + x_2\mathbf{a_2} + \cdots + x_n\mathbf{a_n}$$

矩阵 A 和向量 $\mathbf{x}$ 相乘，**定义为** A 的各列以 $\mathbf{x}$ 对应项为权的线性组合。

> 注意：只在 A 的列数等于 $\mathbf{x}$ 的项数时有定义.

② 矩阵方程

> $matrix\ equation$

回忆：1.3 ⑤ 向量方程 ($vector\ equation$)

> 向量方程 $x_1\mathbf{a_1} + x_2\mathbf{a_2} + \cdots + x_n\mathbf{a_n} = \mathbf{b}$
>
> 和增广矩阵为 $\begin{bmatrix} \mathbf{a_1} & \mathbf{a_2} & \cdots & \mathbf{a_n} & \mathbf{b} \end{bmatrix}$ 的线性方程组有相同的解集。

根据 $A\mathbf{x}$ 定义，把向量方程写成 $A\mathbf{x} = \mathbf{b}$ 的形式，叫做矩阵方程 ($matrix\ equation$)。

其中，$A = \begin{bmatrix} \mathbf{a_1} & \mathbf{a_2} & \cdots & \mathbf{a_n} \end{bmatrix}$，$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$.

③ Theorem 1.3.1

矩阵方程 $A\mathbf{x} = \mathbf{b}$

向量方程 $x_1\mathbf{a_1} + x_2\mathbf{a_2} + \cdots + x_n\mathbf{a_n} = \mathbf{b}$

增广矩阵为 $\begin{bmatrix} \mathbf{a_1} & \mathbf{a_2} & \cdots & \mathbf{a_n} & \mathbf{b} \end{bmatrix}$ 的线性方程组

有相同的解集。

> 这意味着，解 $A\mathbf{x} = \mathbf{b}$，就是解增广矩阵为 $\begin{bmatrix} \mathbf{a_1} & \mathbf{a_2} & \cdots & \mathbf{a_n} & \mathbf{b} \end{bmatrix}$ 的线性方程组，具体解法参见 1.2 ④ ⑤

④ Theorem 1.3.2

Let A be a $m \times n$ matrix. 下面四个命题等价：

(1) For each $\mathbf{b} \in \mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.

(2) Each $\mathbf{b} \in \mathbb{R}^m$ is a linear combination of the columns of A.

(3) The columns of A span $\mathbb{R}^m$.

(4) A has a pivot position in every row.

> $pivot\ position$：见 1.2 ③.
>
> 注意 (4) 中讨论的是**系数矩阵**，不是增广矩阵；
>
> 若**增广矩阵** $\begin{bmatrix} A & \mathbf{b} \end{bmatrix}$ has a pivot position in every row, $A\mathbf{x} = \mathbf{b}$ 可能相容，也可能不相容。

⑤ 单位矩阵

> $identity\ matrix$：主对角线为1，其他项为 0 的方阵（行列数相同）

例：$3 \times 3\ identity\ matrix$

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

> 主对角线（$diagonal$）：左上到右下

$A(\mathbf{u}+\mathbf{v}) = A\mathbf{u} + A\mathbf{v}$

$A(c\mathbf{u}) = c(A\mathbf{u})$

1.5 线性方程组的解集

> SOLUTION SETS OF LINEAR SYSTEMS

homogeneous 齐次的？？？

$A\mathbf{x} = \mathbf{0}$

> $A \in \mathbb{R}^{m \times n},\ \mathbf{x} \in \mathbb{R}^n,\ \mathbf{0} \in \mathbb{R}^m.$

trivial solution 0

> 平凡解

nontrivial solution

parametric vector equation

> x = su + tv

parametric vector form

nonhomogeneous linear system

x = p + tv (t in R)

translate

theorem 6

1.6 线性方程组的应用

> applications of linear systems

略

1.7 线性无关

> linear independence

linear independence

1 linearly independent / linearly dependent

2 linear dependence relation

characterization of linearly dependent sets

theorem 8

theorem 9

1.8 线性变换的介绍

> INTRODUCTION TO LINEAR TRANSFORMATIONS

transformation

domain / codomain

image / range

shear transformation

线性变换

A transformation T is linear if

i) $T(\mathbf{u}+\mathbf{v})=T(\mathbf{u})+T(\mathbf{v})$ for all $\mathbf{u},\ \mathbf{v}$ in the domain of T ;

ii) $T(c\mathbf{u})=cT(\mathbf{u})$ for all scalars c and all $\mathbf{u}$ in the domain of T.

结论：

$T(\mathbf{0})=\mathbf{0}$

> 两个 $\mathbf{0}$ 维度不同

$T(c\mathbf{u}+d\mathbf{v})=cT(\mathbf{u})+dT(\mathbf{v})$

contraction / dilation

1.9 线性变换的矩阵

> THE MATRIX OF A LINEAR TRANSFORMATION

theorem 10

standard matrix for the linear transformation

1.10 线性模型

> LINEAR MODELS IN BUSINESS, SCIENCE, AND ENGINEERING

linear difference equation / recurrence relation

2. 矩阵代数

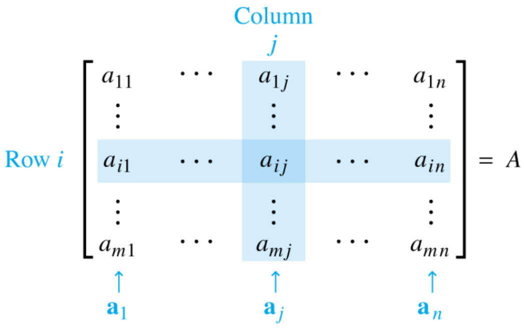
> $Matrix\ Algebra$

2.1 矩阵运算

> $matrix\ operations$

① (i, j) 元素

> (i, j) - $entry$



$m \times n$ $matrix$ A 的各列是 $\mathbb{R}^m$ 中的向量（因为每列有 m 个元素）.

$A = [\,\mathbf{a_1} \quad \mathbf{a_2} \quad \cdots \quad \mathbf{a_n}\,]$

② 对角线元素

> $diagonal\ entries$: $a_{11}, a_{22}, a_{33}, \ \ldots$ （左上到右下）

$$\begin{bmatrix} \color{red}{1} & 1 & 1 \\ 1 & \color{red}{1} & 1 \\ 1 & 1 & \color{red}{1} \end{bmatrix}$$

③ 主对角线

> $main\ diagonal$: 对角线元素组成的线（左上到右下）

④ 对角矩阵

> $diagonal\ matrix$: 非对角线元素全为 0 的**方阵**

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

⑤ 零矩阵

> $zero\ matrix$: 元素全为 0

⑥ 矩阵相等

> $equal$: $size$ 相同，对应 $entry$ 相等

⑦ 和

> sum

设 $A=\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$, $B=\begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix}$, $C=\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$

则 $A+B=\begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & a_{13}+b_{13} \\ a_{21}+b_{21} & a_{22}+b_{22} & a_{23}+b_{23} \end{bmatrix}$, $A+C$ 没有定义（$size$ 不同）

⑧ 标量乘法

> $scalar\ multiple$: 若 r 是标量，A 是矩阵，则标量乘法 rA 是一个矩阵，每一列是 A 对应列的 r 倍.

设 $A=\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$ ，则 $rA=\begin{bmatrix} ra_{11} & ra_{12} & ra_{13} \\ ra_{21} & ra_{22} & ra_{23} \end{bmatrix}$

*定义 $-A$ 为 $(-1)A$, $A-B$ 为 $A+(-1)B$.

⑨ Theorem 2.1

A, B, C 维数相同，r, s 为标量，则

$A+B=B+A$

$(A+B)+C=A+(B+C)$

$A+0=A$

$r(A+B)=rA+rB$

$(r+s)A=rA+sA$

$r(sA)=(rs)A$

> 易得，左右对应项相等

⑩矩阵乘法

我们希望把复合映射 $A(B\mathbf{x})$ （先乘 B 再乘A）表示为乘以一个矩阵的变换. 此矩阵记为 AB , 即

$A(B\mathbf{x})=(AB)\mathbf{x}$

若 A 是 $m\times n$ 矩阵，B 是 $n\times p$ 矩阵，$\mathbf{x}\in\mathbb{R}^p$, 则

$$B\mathbf{x}=\begin{bmatrix} \mathbf{b_1} & \mathbf{b_1} & \cdots & \mathbf{b_p} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}=x_1\mathbf{b_1}+x_2\mathbf{b_2}+\cdots+x_p\mathbf{b_p}$$

$A(B\mathbf{x}) = A(x_1\mathbf{b_1} + x_2\mathbf{b_2} + \cdots + x_p\mathbf{b_p}) = x_1A\mathbf{b_1} + x_2A\mathbf{b_2} + \cdots + x_pA\mathbf{b_p}$

向量 $A(B\mathbf{x})$ 是向量 $A\mathbf{b_1}, \cdots, A\mathbf{b_p}$ 的线性组合，以 $\mathbf{x}$ 的 $entry$ 为权.

$$A(B\mathbf{x}) = [\, A\mathbf{b_1} \quad A\mathbf{b_2} \quad \cdots \quad A\mathbf{b_p} \,] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}$$

矩阵乘积的列定义：若 A 是 $m \times n$ 矩阵，B 是 $n \times p$ 矩阵，B 的列是 $\mathbf{b_1}, \cdots, \mathbf{b_p}$，则乘积 AB 是 $m \times p$ 矩阵，且

$$AB = [\, A\mathbf{b_1} \quad A\mathbf{b_2} \quad \cdots \quad A\mathbf{b_p} \,]$$

> AB 的列是以 B 对应列的元素为权，对 A 各列进行线性组合出来的（AB 每列 m 个元素，和 A 一致；共 n 列，和 B 一致）

A 的列数必须等于 B 的行数（才能使线性组合有定义）

⑪ 行列法则

> $Row - column\ rule\ for\ computing\ AB$

$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{n1}b_{nj}$

> AB 的第 i 行第 j 列元素是 A 的第 i 行与 B 的第 j 列对应元素乘积之和。

记 $row_i(A)$ 表示矩阵 A 的第 i 行，则

$row_i(AB) = row_i(A) \cdot B$

⑫ Theorem 2.2

>

$A(BC) = (AB)C$

$A(B + C) = AB + AC$

$(B + C)A = BA + CA$

$r(AB) = (rA)B = A(rB)$

$I_mA = A = AI_n$

易错点 1

通常来说，$AB \neq BA$.

> 这是 $matrix\ algebra$ 和实数代数的关键区别。
>
> 必要时可强调 AB 的相对位置：$A\ is\ right - multiplied\ by\ B$ （A 被 B 右乘）/ $B\ is\ left - multiplied\ by\ A$ （B 被 A 左乘）

> The cancellation laws do not hold for matrix multiplication. That is, if $AB = AC$, then it is not true in general that $B = C$.
>
> If a product AB is the zero matrix, you cannot conclude in general that either $A = 0$ or $B = 0$.

A^k / A^0

transpose

A^T

Theorem 3

$(A^T)^T = A$

$(A+B)^T = A^T + B^T$

$(rA)^T = rA^T$

$(AB)^T = B^T A^T$

> 注意方向调转了

2.2 矩阵的逆

> THE INVERSE OF A MATRIX

invertible

可逆矩阵一定是方阵

$A^{-1}A = AA^{-1} = I$

Theorem 4 二阶方阵可逆性判别 & 逆矩阵求法

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If $ad - bc \neq 0$, then A is invertible and

$A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

If $ad - bc = 0$, then A is not invertible.

determinant

Theorem 6

$(A^{-1})^{-1} = A$

$(AB)^{-1} = B^{-1}A^{-1}$

$(A^T)^{-1} = (A^{-1})^T$

Theorem 7 高阶方阵可逆性判别 & 逆矩阵求法（初等变换法）

2.3 可逆矩阵的特征

CHARACTERIZATIONS OF INVERTIBLE MATRICES

Theorem 8

Let A be a square $n \times n$ matrix.

all true or all false

① A is an invertible matrix.

② A is row equivalent to the $n \times n$ identity matrix.

③ A has n pivot position.

④ The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. (for any $\mathbf{x} \neq \mathbf{0}$, $A\mathbf{x} \neq \mathbf{0}$.)

⑤ The columns of A form a linearly independent set.

⑥ The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.

⑦ The equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$.

⑧ The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^n$.

⑨ There is an $n \times n$ matrix C such that $CA = I$.

⑩ There is an $n \times n$ matrix D such that $AD = I$.

⑪ A^T is an invertible matrix.

> applies only to square matrices

A linear transformation "invertible"

A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is said to be **invertible** if there exists a function $S : \mathbb{R}^n \to \mathbb{R}^n$ such that

$S(T(\mathbf{x})) = \mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$.   (1)

$T(S(\mathbf{x})) = \mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$. (2)

Theorem 9

Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation and let A be the standard matrix for T. Then T is invertible if and only if A is an invertible matrix. In that case, the linear transformation S given by $S(\mathbf{x}) = A^{-1}\mathbf{x}$ is the unique function satisfying equation (1) and (2).

2.4 分块矩阵

> $partitioned\ matrices$

① 分块矩阵

blocks / submatrices

partitioned / block matrix

② 加法与标量乘法

③ 分块矩阵的乘法

④ Theorem 10

Column—Row Expansion of AB AB的列行展开

⑤ 分块矩阵的逆

block upper triangular 分块上三角矩阵

⑥ 分块对角矩阵

> $block\ diagonal\ matrix$

2.5 矩阵因式分解

> MATRIX FACTORIZATIONS

① LU 分解

LU factorization

② LU 分解算法

> Algorithm for an LU Factorization

2.6 列昂惕夫投入产出模型

> THE LEONTIEF INPUTOUTPUT MODEL

> Leontief Input-Output Model（勒庞蒂夫投入产出模型）是一种经济分析工具，用于研究不同产业之间的相互依赖关系和经济活动的影响。
>
> 该模型是由美国经济学家沃西里·勒庞蒂夫（Wassily Leontief）在20世纪40年代开发的，他因此获得了1973年的诺贝尔经济学奖。该模型基于一个假设，即一个经济系统可以被划分为不同的产业部门，这些部门之间存在着复杂的物质和货币交换关系。
>
> Leontief模型的核心思想是，不同的产业部门之间存在着输入和输出的关系。每个产业部门都需要使用其他产业部门的产品作为投入，同时也会向其他产业部门提供产品作为产出。通过分析这些输入和输出关系，可以揭示不同部门之间的相互依赖性，以及经济活动的波及效应。
>
> 通过使用Leontief模型，经济学家可以评估不同产业部门之间的关系，以及某个部门的变化如何影响整个经济系统。这种模型可以用于预测和评估政策变化、经济冲击和其他因素对经济的影响。此外，Leontief模型还可以帮助分析国际贸易、资源分配和产业结构等问题。
>
> 总而言之，Leontief Input-Output Model是一种用于分析经济系统中产业部门相互依赖关系的模型，它提供了一种量化的方法来研究经济活动的影响和波及效应。

production vector / final demand vector

unit consumption vector

Theorem 11

2.7 计算机图形学中的应用

Applications to Computer Graphics

略

2.8 $\mathbb{R}^n$ 的子空间

> SUBSPACES OF $\mathbb{R}^n$

Definition: A subspace of $\mathbb{R}^n$ is any set H in $\mathbb{R}^n$ that has three properties:

a) The zero vector is in H.

b) For each $\mathbf{u}$ and $\mathbf{v}$ in H, the sum $\mathbf{u}+\mathbf{v}$ is in H .

c) For each $\mathbf{u}$ in H and each scalar c, the vector $c\mathbf{u}$ is in H.

column space

null space

Theorem 12

basis

standard basis

Theorem 13

2.9 维数与秩

> DIMENSION AND RANK

coordinates of x relative to the basis

coordinate vector of x

dimension

rank

Theorem 14

Theorem 15

The Invertible Theorem (continued)

> 目测这条 theorem 应该是整章的核心

3. 行列式

> determinants

4. 向量空间

> Vector Spaces

vector space

subspace

zero subspace

Theorem 1

null space

An Explicit Description of NulA

basis

standard basis

The pivot columns of a matrix A form a basis for Col A.

THE UNIQUE REPRESENTATION THEOREM

coordinate vector of x

coordinate mapping

change-of-coordinates matrix

isomorphism 同构

THE DIMENSION OF A VECTOR SPACE

finite-dimensional

infinite-dimensional

5. 特征值与特征向量

Eigenvalues and Eigenvectors

Theorem 5 是重点

6. 正交性和最小二乘法

> $Orthogonality\ and\ Least\ Squares$

6.1 内积，长度和正交性

> $inner\ product,\ length,\ and\ orthogonality$

① 内积的定义

若 $\mathbf{u}$，$\mathbf{v}$ 是 $\mathbb{R}^n$ 中的向量，可视为 $n \times 1$ 矩阵.

定义 $\mathbf{u}^T\mathbf{v}$ 为 $\mathbf{u}$ 和 $\mathbf{v}$ 的 $inner\ product$，记为 $\mathbf{u} \cdot \mathbf{v}$.

> 是一个标量，不用加括号

本质是用矩阵乘法的形式重新定义了向量点乘。

注意需要两个列向量，列数超过 1 的没有定义。

② 内积的性质

③ 向量的长度（范数）

和向量模长定义一样

④ 向量归一化

> $normalize$

把一个向量除以它的模长，得到同方向上的单位向量。

⑤ 两个向量的距离

> $distance$

$dist(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$

⑥ 两个向量正交

> $Two\ vectors\ are\ orthogonal\ to\ each\ other$

$\mathbf{u} \cdot \mathbf{v} = 0$

零向量和任意向量正交。

定理：Two vectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal if and only if

$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$

⑦ 正交补

> $Orthogonal\ Complements$

向量 $\mathbf{z}$ 正交于子空间 W 中的每一个向量，则向量 $\mathbf{z}$ 与 W 正交。

所有 $\mathbf{z}$ 构成的集合为 W 的正交补，记作 $W^\perp$，读作 "$W\ perpendicular$" 或 "$W\ perp$".

定理：

$(Row\ A)^\perp = Nul\ A$
$(Col\ A)^\perp = Nul\ A^T$

⑧ $\mathbb{R}^2$ 和 $\mathbb{R}^3$ 中的角度

$cos\ \theta = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}$

高于三维，直接定义. 统计学中称为两个向量的相关系数（$correlation\ coefficient$）.

6.2 正交集

> $orthogonal\ sets$

① 正交集的定义

$\mathbb{R}^n$ 中的向量集 $\{\mathbf{u_1},\ \mathbf{u_2},\ \cdots,\ \mathbf{u_p}\}$ 称为正交集，若其中的向量两两正交。

定理：

$S = \{\mathbf{u_1},\ \mathbf{u_2},\ \cdots,\ \mathbf{u_p}\}$ 是 $\mathbb{R}^n$ 中的非零向量构成的正交集，则 S 线性独立，因此是 S 张成的子空间的一组 $basis.$

② 正交基的定义

$\mathbb{R}^n$ 中子空间 W 的一个正交基，既是 W 的一个基，也是正交集。

> 就是 $basis$ 中正交的那些。

③ 单位正交集

> $orthonormal\ set$

单位向量构成的正交集.

④ 单位正交基

> $orthonormal\ basis$

W 为单位正交集张成的子空间，则该单位正交集也叫做 W 的单位正交基。

> 单位正交集显然线性独立

注意：$basis$ 范围比 set 更窄，$basis$ 需要有对应的张成空间，并且要满足线性独立.

定理：

An $m \times n$ matrix U has orthonormal columns if and only if $U^TU = I$.

定理：

Let U be an $m \times n$ matrix with orthonormal columns, and let $\mathbf{x}$ and $\mathbf{y}$ be in $\mathbb{R}^n$.

Then,

$$\|U\mathbf{x}\| = \|\mathbf{x}\|$$
$$(U\mathbf{x})\cdot(U\mathbf{y}) = \mathbf{x}\cdot\mathbf{y}$$
$$(U\mathbf{x})\cdot(U\mathbf{y}) = 0 \Leftrightarrow \mathbf{x}\cdot\mathbf{y} = 0$$

> 第一条和第三条说明：有单位正交列的矩阵，对应的线性映射 $\mathbf{x} \longmapsto U\mathbf{x}$，保留了长度和正交性.

6.3 正交投影

> $orthogonal\ projections$

① 正交分解定理

Let W be a subspace of $\mathbb{R}^n$. Then each $\mathbf{y}$ in $\mathbb{R}^n$ can be written uniquely in the form

$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$

where $\hat{\mathbf{y}}$ is in W and $\mathbf{z}$ is in $W^\perp$.

> $\hat{\mathbf{y}}$ 就是投影向量

这条定理包含两条信息.

一是，If $\{\mathbf{u_1},\ \mathbf{u_2},\ \cdots,\ \mathbf{u_p}\}$ is an orthogonal basis of W,

$\mathbf{z} = \mathbf{y} - \frac{\mathbf{y}\cdot\mathbf{u_1}}{\mathbf{u_1}\cdot\mathbf{u_1}}\mathbf{u_1} + \cdots + \frac{\mathbf{y}\cdot\mathbf{u_p}}{\mathbf{u_p}\cdot\mathbf{u_p}}\mathbf{u_p}$ 和 W 中的任意向量正交.

二是，对于给定 W，这种正交分解是唯一的（即分解成一个在 W 里，一个在 $W^\perp$ 里，那么分量一定是 $\hat{\mathbf{y}}$ 和 $\mathbf{z}$.）

> 分解的结果只取决于 W 本身，而与 W 正交基的选取无关. 选不同的正交基可以算出同一个结果.

$\hat{\mathbf{y}}$ 的求法：

If $\{\mathbf{u_1},\ \mathbf{u_2},\ \cdots,\ \mathbf{u_p}\}$ is an orthogonal basis of W, then

$\hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u_1}}{\mathbf{u_1}\cdot\mathbf{u_1}}\mathbf{u_1} + \cdots + \frac{\mathbf{y}\cdot\mathbf{u_p}}{\mathbf{u_p}\cdot\mathbf{u_p}}\mathbf{u_p}$

$\mathbf{z}$ 的求法：

先求 $\hat{\mathbf{y}}$，然后 $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$.

验算方法：

计算 $\mathbf{z}\cdot\mathbf{u_1}$，若为 0，则计算无误.

② 正交投影

上述 $\hat{\mathbf{y}}$ 称为 $\mathbf{y}$ 在 W 上的正交投影，记作 $proj_W\mathbf{y}$.

> $orthogonal\ projection\ of\ \mathbf{y}\ onto\ W$

$\mathbf{y} \in W \Rightarrow proj_W\mathbf{y} = \mathbf{y}$

最佳逼近定理：

Let W be a subspace of $\mathbb{R}^n$, let $\mathbf{y}$ be any vector in $\mathbb{R}^n$, and let $\hat{\mathbf{y}}$ be the orthogonal projection of $\mathbf{y}$ onto W. Then $\hat{\mathbf{y}}$ is the closest point in W to $\mathbf{y}$, in the sense that

$\|\mathbf{y} - \hat{\mathbf{y}}\| < \|\mathbf{y} - \mathbf{v}\|$

for all $\mathbf{v}$ in W distinct from $\hat{\mathbf{y}}$.

向量 $\hat{\mathbf{y}}$ 称为 W 中元素对 $\mathbf{y}$ 的最佳逼近.

> *the best approximation to* $\mathbf{y}$ *by elements of* W.

$\|\mathbf{y}-\mathbf{v}\|$ 称为误差，$\mathbf{v}=\hat{\mathbf{y}}$ 时取到最小值.

简化定理：

我们已知，

If $\{\mathbf{u_1},\ \mathbf{u_2},\ \cdots,\ \mathbf{u_p}\}$ is an *orthogonal* basis of W, then

$proj_W\mathbf{y}=\hat{\mathbf{y}}=\frac{\mathbf{y}\cdot\mathbf{u_1}}{\mathbf{u_1}\cdot\mathbf{u_1}}\mathbf{u_1}+\cdots+\frac{\mathbf{y}\cdot\mathbf{u_p}}{\mathbf{u_p}\cdot\mathbf{u_p}}\mathbf{u_p}$

进一步得到，

If $\{\mathbf{u_1},\ \mathbf{u_2},\ \cdots,\ \mathbf{u_p}\}$ is an *orthonormal* basis of W, then

$$\begin{aligned} proj_W\mathbf{y}=\hat{\mathbf{y}}&=\frac{\mathbf{y}\cdot\mathbf{u_1}}{1}\mathbf{u_1}+\cdots+\frac{\mathbf{y}\cdot\mathbf{u_p}}{1}\mathbf{u_p}\\ &=(\mathbf{y}\cdot\mathbf{u_1})\mathbf{u_1}+(\mathbf{y}\cdot\mathbf{u_2})\mathbf{u_2}+\cdots+(\mathbf{y}\cdot\mathbf{u_p})\mathbf{u_p}\\ &=(\mathbf{u_1^T y})\mathbf{u_1}+\cdots+(\mathbf{u_p^T y})\mathbf{u_p}\\ &=UU^T\mathbf{y}\end{aligned}$$

where $U=[\,\mathbf{u_1}\quad\mathbf{u_2}\quad\cdots\quad\mathbf{u_p}\,]$.

6.4 Gram-Schmidt 方法

> *the Gram − Schmidt process*

用途：构造 $\mathbb{R}^n$ 中任何非零子空间的正交基或标准正交基.

① Gram-Schmidt 方法

Given a basis $\{\mathbf{x_1},\ \cdots,\ \mathbf{x_p}\}$ for a nonzero subspace W of $\mathbb{R}^n$, define

$$\begin{aligned} \mathbf{v_1}&=\mathbf{x_1}\\ \mathbf{v_2}&=\mathbf{x_2}-\frac{\mathbf{x_2}\cdot\mathbf{v_1}}{\mathbf{v_1}\cdot\mathbf{v_1}}\mathbf{v_1}\\ \mathbf{v_3}&=\mathbf{x_3}-\frac{\mathbf{x_3}\cdot\mathbf{v_1}}{\mathbf{v_1}\cdot\mathbf{v_1}}\mathbf{v_1}-\frac{\mathbf{x_3}\cdot\mathbf{v_2}}{\mathbf{v_2}\cdot\mathbf{v_2}}\mathbf{v_2}\\ &\vdots\\ \mathbf{v_p}&=\mathbf{x_p}-\frac{\mathbf{x_p}\cdot\mathbf{v_1}}{\mathbf{v_1}\cdot\mathbf{v_1}}\mathbf{v_1}-\frac{\mathbf{x_p}\cdot\mathbf{v_2}}{\mathbf{v_2}\cdot\mathbf{v_2}}\mathbf{v_2}-\cdots-\frac{\mathbf{x_p}\cdot\mathbf{v_{p-1}}}{\mathbf{v_{p-1}}\cdot\mathbf{v_{p-1}}}\mathbf{v_{p-1}}\end{aligned}$$

结论1：$\{\mathbf{v_1},\ \cdots,\ \mathbf{v_p}\}$ is an orthogonal basis for W.

结论2：$Span\,\{\mathbf{v_1},\ \cdots,\ \mathbf{v_k}\}=Span\,\{\mathbf{x_1},\ \cdots,\ \mathbf{x_k}\}$ for $1\le k\le p$.

> 注意，前提是 $\{x_1,\ \cdots,\ x_p\}$ 已经是一组 $basis$，即它们线性独立

② 矩阵 QR 分解

If A is an $m \times n$ matrix with linearly independent columns, then A can be factored as $A = QR$, where Q is an $m \times n$ matrix whose columns form an orthonormal basis for $Col\ A$ and R is an $n \times n$ upper triangular invertible matrix with positive entries on its diagonal.

> A 要满足各列线性独立，原因是证明中要用到 A 的各列形成 $Col\ A$ 的 $basis$（即各列独立）.
>
> Q 的构造：以 A 的各列为 $basis$，用 $Gram-Schmidt$ 方法生成新单位正交基的各元素作为 Q 的列.
>
> 注意用 $Gram-Schmidt$ 方法生成的正交基可能未归一化，需要归一化才是标准的 Q.

为什么要大费周章地把 Q 归一化呢？原因在于 R 的构造.

引理：

An $m \times n$ matrix U has orthonormal columns if and only if $U^T U = I$.

当 Q 的各列是单位正交的，有 $Q^T Q = I$.

$Q^T A = Q^T (QR) = R$.

只要把得到的 Q 的转置乘到 A 上，就得到了 R.

6.5 最小二乘问题

> $least-squares\ problems$

① 定义

If A is $m \times n$ and $\mathbf{b}$ is in $\mathbb{R}^m$, a least-squares solution of $A\mathbf{x} = \mathbf{b}$ is an $\hat{\mathbf{x}}$ in $\mathbb{R}^n$ such that

$\|\mathbf{b} - A\hat{\mathbf{x}}\| \leq \|\mathbf{b} - A\mathbf{x}\|$

> $A\hat{\mathbf{x}}$ 是 $Col\ A$ 中离 $\mathbf{b}$ 最近的点.
>
> 能解则解，不能解则逼近.

② 解法

6.6 线性模型中的应用

略

6.7 内积空间

> $inner\ product\ spaces$

6.8 内积空间的应用

略

7. 对称矩阵和二次型

Symmetric Matrices and Quadratic Forms

对称矩阵

谱分解

奇异值分解

8. 向量空间的几何学

> The Geometry of Vector Spaces

9. Optimization

10. Finite - State Marcov Chains

Appendixes

Weekly Exercise

exercise 1

Group 1 : 2 4 5 7 10

1. (Lay 1.1, Question 26) Construct three augmented matrices for linear systems whose solution set is $x_1 = -2\ ,\ x_2 = 1\ ,\ x_3 = 0$

$\begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ or $\begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & -2 \end{bmatrix}$ etc.

2. (Lay 1.1, Question 28) Suppose $a,\ b,\ c$, and d are constants such that a is not zero and the system below is consistent for all possible values of f and g.

$$\begin{aligned} ax_1 &+ bx_2 &= f \\ cx_1 &+ dx_2 &= g \end{aligned}$$

What can you say about the numbers $a,\ b,\ c,$ and d ? Justify your answer.

> 关键是 a 非零的使用
> 和 f g 任取

3. (Lay 1.2, Question 21) Mark each statement True or False. Justify your answer.

(a) In some cases, a matrix may be row reduced to more than one matrix in reduced echelon form, using different sequences of row operations.

(b) The row reduction algorithm applies only to augmented matrices for a linear system.

(c) A basic variable in a linear system is a variable that corresponds to a pivot column in the coefficient matrix.

(d) Finding a parametric description of the solution set of a linear system is the same as solving the system.

(e) If one row in an echelon form of an augmented matrix is [0 0 0 5 0], then the associated linear system is inconsistent.

(a)F

> echelon form 有无穷个，reduced echelon form 只有一个（Theorem 1）

反证法：suppose A 化简成两个不同的 reduced echelon form $f\ and\ g$

$$\begin{matrix} & & F_i \\ \end{matrix}$$
$$\begin{bmatrix} 1 & \cdots & x \\ 0 & & \\ \vdots & & \\ 0 & & \end{bmatrix}$$

从左到右，设第 i 列开始不同

有三种情况

① fi gi both pivot column

② both not

③ one of them is pivot column

> 若 s 行 i 列 要从 0 转成 1，则 s - 尾 行， 1 ~ i -1 列 这块区域不可能相同（不能通过行变换得到）

三种都不可能

(b) F

> elementary operation 是对任意矩阵适用的，不分是否 augmented

(c) $True$. Basic variables are defined after equation (4).

F / T ? $True$. This statement is at the beginning of Parametric Descriptions of Solution Sets.

F

4. (Lay 1.2, Question 24) Suppose a system of linear equations has 3 x 5 augmented matrix whose fifth column is a pivot column. Is the system consistent? Why (or why not)?

   它的 reduced echelon form 一定有一行 0 0 0 0 1 （inconsistent）

5. (Lay 1.3, Question 33) Use the vectors **u** = $(u_1, \ ..., \ u_n)$, **v** = $(v_1, \ ..., \ v_n)$, and **w** = $(w_1, \ ..., \ w_n)$ to verify the following algebraic properties of R^n .

   (a) (**u** + **v**) + **w** = **u** + (**v** + **w**)

   (b) c(**u** + v) = c**u** + c**v** for each scalar c

> 可以讨论第 i 项相同，然后 i 遍历 1 到 n

6. (Lay 1.3, Question 34) Use the vector **u** = ($u_1, \ ..., \ u_n$) to verify the following algebraic properties of R^n.

   (a) **u** + (-**u**) = **0**
   (b) c(d**u**) = (cd)**u** for all scalars c and d

> 同上

7. (Lay 1.4 Question 35) Let A be a 3 x 4 matrix, let y_1 and y_2 be vectors in R^3 , and let **w** = y_1 + y_2 . Suppose y_1 = Ax_1 and y_2 = A x_2 for some vectors x_1 and x_2 in R^4 . What fact allows you to conclude that the system A**x** = **w** is consistent? (Note x_1 and x_2 denotes vectors, not scalar entries in vectors.)

8. (Lay 1.4, Question 36) Let A be a 5 x 3 matrix, let **y** be a vector in R^3 , and let **z** be a vector in R^5 . Suppose A**y** = **z**. What fact allows you to conclude that the system A**x** = 4**z** is consistent?

   > 至少有一解 $\mathbf{x} = 4\mathbf{y}$

9. (Lay 1.5, Question 37) Constract a 2×2 matrix A such that the solution set of the equation A**x** = **0** is the line in R^2 through (4,1) and the origin. Then, find a vector **b** in R^2 such that the solution set of A**x** = **b** is not a line in R^2 parallel to the solution set of $A\mathbf{x} = \mathbf{0}$. Why does this not contradict to the following theorem?

   Theorem. Suppose the equation A**x** = **b** is consistent for some given **b**, and let **p** be a solution. Then the solution set of A**x** = **b** is the set of all vectors of the form **w** = **p** + v_h , where v_h is any solution of the homogeneous equation A**x** = **0**.

   > Theorem 的假设是有解，我们可以取b 使 它无解（没有直线，就不会平行）。

   ? ? ?

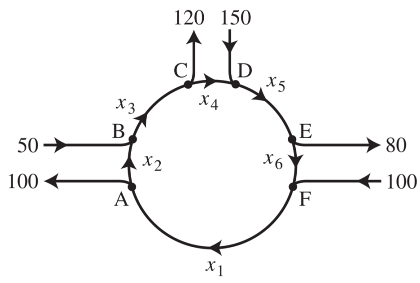
   $$A = \begin{bmatrix} 1 & -4 \\ 1 & -4 \end{bmatrix}$$

   $$b = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

   b的两个项不同即可使x无解

10. (Lay 1.5, Question 40) Let A be an m x n matrix, let u and v be vectors in R^n with the property that Au = 0 and Av = 0. Explain why A(u + v) must be zero vector. Then explain why A(cu + dv) = 0 for each pair of scalars c and d.

11. (Lay 1.6 Question 14) Intersections in England are often constructed as one-way "roundabouts" such as the one shown in the figure. Assume that traffic must travel in the directions shown. Find the general solution of the network flow. Find the smallest possible value for x_6.



列方程组

70

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & -1 & 100 \\ \ldots \end{bmatrix}$$

12. (Lay 1.7, Question 36) If v_1 , ... , v_4 are in R^4 and v_3 is not a linear combination of $v_1,\ v_2,$ and v_4 , is $\{v_1,\ v_2,\ v_3,\ v_4\}$ linearly independent? Prove if you think so or give a counterexample otherwise.

都有可能

independent if and only $a\mathbf{V}_1 + b\mathbf{V}_2 + c\mathbf{V}_3 + d\mathbf{V}_4 = \mathbf{0}$ 的解为 $a = b = c = d = 0$

> 可以取一些零向量使 $a/b/c/d$ 不为 0

Group 1 work

2. (Lay 1.1, Question 28) Suppose $a,\ b,\ c$, and d are constants such that a is not zero and the system below is consistent for all possible values of f and g.

$$\begin{aligned} ax_1 &+ bx_2 &= f \\ cx_1 &+ dx_2 &= g \end{aligned}$$

What can you say about the numbers $a,\ b,\ c,$ and d ? Justify your answer.

Solution:

The augmented matrix for the system is

$$\begin{bmatrix} a & b & f \\ c & d & g \end{bmatrix}$$

Row reduced it. Because a is not zero,

$$\begin{bmatrix} a & b & f \\ c & d & g \end{bmatrix} \sim \begin{bmatrix} a & b & f \\ 0 & d - \frac{bc}{a} & g - \frac{fc}{a} \end{bmatrix}$$

If $d - \frac{bc}{a} = 0$:

Because f and g are arbitrary, there exist some f and g that $g - \frac{fc}{a} \neq 0$, which makes the system inconsistent (by **Existence and Uniqueness Theorem**) and contradicts the question.

Hence

$$d-\frac{bc}{a}\neq 0\Rightarrow bc\neq ad$$

Appendix ① 充分性

> 用必要性足以完成题目（consistent for arbitrary f and $g\Rightarrow bc\neq ad$）；

下面给出充分性证明（$bc\neq ad\Rightarrow$ consistent for arbitrary f and g）：

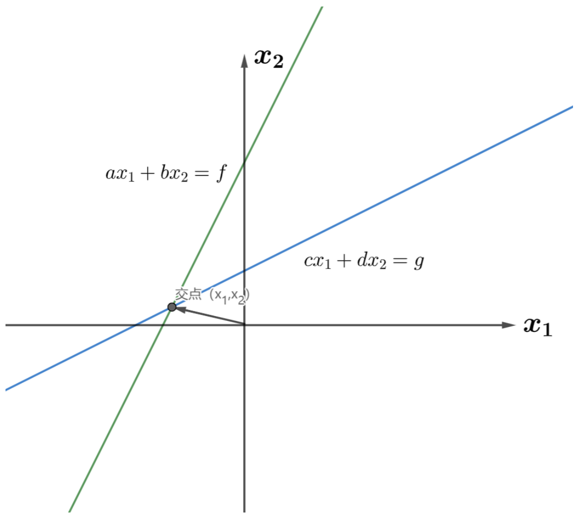
When $bc\neq ad$,

$$\begin{bmatrix} a & b & f \\ 0 & d-\frac{bc}{a} & g-\frac{fc}{a} \end{bmatrix}\sim\begin{bmatrix} a & b & f \\ 0 & 1 & \frac{g-\frac{fc}{a}}{d-\frac{bc}{a}} \end{bmatrix}\sim\begin{bmatrix} a & 0 & f-b\cdot\frac{g-\frac{fc}{a}}{d-\frac{bc}{a}} \\ 0 & 1 & \frac{g-\frac{fc}{a}}{d-\frac{bc}{a}} \end{bmatrix}\sim\begin{bmatrix} 1 & 0 & \frac{f-b\cdot\frac{g-\frac{fc}{a}}{d-\frac{bc}{a}}}{a} \\ 0 & 1 & \frac{g-\frac{fc}{a}}{d-\frac{bc}{a}} \end{bmatrix}$$

The system is consistent for all possible values of f and g.

More precisely, it has and only has one solution $\vec{x}=\begin{bmatrix} \frac{f-b\cdot\frac{g-\frac{fc}{a}}{d-\frac{bc}{a}}}{a} \\ \frac{g-\frac{fc}{a}}{d-\frac{bc}{a}} \end{bmatrix}$

Appendix ② Geometric Description



4. (Lay 1.2, Question 24) Suppose a system of linear equations has a 3×5 augmented matrix whose fifth column is a pivot column. Is the system consistent? Why (or why not)?

Solution:

The system is **inconsistent**.

Because the fifth column is the rightmost column of the 3×5 augmented matrix $\begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix}$. By Theorem 2, If the fifth column is a pivot column, the echelon form of the original matrix will have one row $\begin{bmatrix} 0 & 0 & 0 & 0 & b \end{bmatrix}$ with b nonzero, which results in the system having no solution (inconsistent).

5. (Lay 1.3, Question 33) Use the vector $\mathbf{u} = (u_1, \ldots, u_n)$, $\mathbf{v} = (v_1, \ldots, v_n)$, and $\mathbf{w} = (w_1, \ldots, w_n)$ to verify the following algebraic properties of $\mathbb{R}^n$.
   (a) $(\mathbf{u}+\mathbf{v})+\mathbf{w} = \mathbf{u}+(\mathbf{v}+\mathbf{w})$
   (b) $c(\mathbf{u}+\mathbf{v}) = c\mathbf{u}+c\mathbf{v}$ for each scalar c

Solution:

(a)

$$Left: (\mathbf{u}+\mathbf{v})+\mathbf{w}$$

$$= (\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}) + \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$$

$$= \begin{bmatrix} (u_1+v_1)+w_1 \\ \vdots \\ (u_n+v_n)+w_n \end{bmatrix}$$

$$Right: \mathbf{u}+(\mathbf{v}+\mathbf{w})$$

$$= \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + (\begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} + \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix})$$

$$= \begin{bmatrix} u_1+(v_1+w_1) \\ \vdots \\ u_n+(v_n+w_n) \end{bmatrix}$$

By associativity of addition and definition of equality of vectors ,

$$Left = Right \Rightarrow (\mathbf{u}+\mathbf{v})+\mathbf{w} = \mathbf{u}+(\mathbf{v}+\mathbf{w})$$

(b)

For each scalar c ,

$$Left: c\,(\mathbf{u}+\mathbf{v})$$

$$= c\,(\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix})$$

$$= \begin{bmatrix} c(u_1+v_1) \\ \vdots \\ c(u_n+v_n) \end{bmatrix}$$

$$Right: c\mathbf{u}+c\mathbf{v}$$

$$= c\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + c\begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

$$= \begin{bmatrix} cu_1+cv_1 \\ \vdots \\ cu_n+cv_n \end{bmatrix}$$

By the distributive law and definition of equality of vectors ,

$$Left = Right \Rightarrow c(\mathbf{u}+\mathbf{v}) = c\mathbf{u}+c\mathbf{v}$$

7. (Lay 1.4, Question 35) Let A be a 3×4 matrix, let $\mathbf{y}_1$ and $\mathbf{y}_2$ be vectors in $\mathbb{R}^3$, and let $\mathbf{w} = \mathbf{y}_1 + \mathbf{y}_2$. Suppose $\mathbf{y}_1 = A\mathbf{x}_1$ and $\mathbf{y}_2 = A\mathbf{x}_2$ for some vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ in $\mathbb{R}^4$. What fact allows you to conclude that the system $A\mathbf{x} = \mathbf{w}$ is consistent? (Note: $\mathbf{x}_1$ and $\mathbf{x}_2$ denotes vectors, not scalar entries in vectors.)

Solution:

$$\begin{aligned}\mathbf{w} &= \mathbf{y}_1 + \mathbf{y}_2 \\ &= A\mathbf{x}_1 + A\mathbf{x}_2 \\ &= A(\mathbf{x}_1 + \mathbf{x}_2) \qquad (By\,Theorem\;5)\end{aligned}$$

Hence, the system

$A\mathbf{x} = \mathbf{w}$

has at least one solution $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$ (consistent).

10. (Lay 1.5, Question 40) Let A be an $m \times n$ matrix, let $\mathbf{u}$ and $\mathbf{v}$ be vectors in $\mathbb{R}^n$ with the property that $A\mathbf{u} = \mathbf{0}$ and $A\mathbf{v} = \mathbf{0}$. Explain why $A(\mathbf{u} + \mathbf{v})$ must be the zero vector. Then explain why $A(c\mathbf{u} + d\mathbf{v}) = \mathbf{0}$ for each pair of scalars c and d.

Solution:

Suppose

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}, \mathbf{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}, \mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

Then

$$A(\mathbf{u} + \mathbf{v})$$

$$= A(\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix})$$

$$= \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} u_1 + v_1 \\ \vdots \\ u_n + v_n \end{bmatrix}$$

$$= (u_1 + v_1) \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \cdots + (u_n + v_n) \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

$$= (u_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \cdots + u_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}) + (v_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \cdots + v_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix})$$

$$= \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

$$= A\mathbf{u} + A\mathbf{v}$$

$$= \mathbf{0} + \mathbf{0}$$

$$= \mathbf{0}$$

for each pair of scalars c and d ,

$$A(c\mathbf{u} + d\mathbf{v})$$

$$= A(c \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + d \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix})$$

$$= \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} cu_1 + dv_1 \\ \vdots \\ cu_n + dv_n \end{bmatrix}$$

$$= (cu_1 + dv_1) \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \cdots + (cu_n + dv_n) \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

$$= c\,(u_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \cdots + u_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}) + d\,(v_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \cdots + v_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix})$$

$$= c\,(\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}) + d\,(\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix})$$

$$= c\,(A\mathbf{u}) + d\,(A\mathbf{v})$$

$$= c\,\mathbf{0} + d\,\mathbf{0}$$

$$= \mathbf{0}$$

exercise 2

Group 1 : 1 2 3 4 5

1. Suppose $T : \mathbb{Z}^3 \to \mathbb{Z}^3$ is a linear transformation, where $\mathbb{Z}$ is the set of integers. You are given that $T(\mathbf{x}) = \mathbf{y},\ T(\mathbf{y}) = \mathbf{z}$, and $T(\mathbf{z}) = \mathbf{x} + \mathbf{y}$ for certain $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{Z}^3$, where $\mathbf{x} \neq \mathbf{0}$. Prove that $\mathbf{x}, \mathbf{y}$, and $\mathbf{z}$ are linearly independent. You can use Rational root theorem without proof.

Theorem (Rational root theorem). *If a polynomial $p(x) = a_n x^n + a_{n-1} x^{n-1} + a_{n-2} x^{n-2} + \cdots + a_1 x + a_0$ has rational roots, those roots will be limited to $\pm\frac{p}{q}$ where p is a factor of a_0 and q is a factor of a_n .*

2. Can you find a linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^2$ so that we have $T((1, 0, 3)^T) = (1, 1)^T$ and $T((-2, 0, -6)^T) = (2, 1)^T$? Give an example if you think so or prove the impossibility otherwise.

3. (Lay 1.10, Question 9) In a certain region, about 7% of a city's population moves to the surrounding suburbs each year, and about 5% of the suburban population moves into the city. In 2015, there were 800,000 residents in the city and 500,000 in the suburbs. Set up a difference equation that describes this situation, where $\mathbf{x}_0$ is the initial population in 2015. Then estimate the populations in the city and in the suburbs two years later, in 2017. (Ignore other factors that might influence the population sizes.)

马尔科夫链？画出二维坐标（city，suburb）最后回回归到一个点，满足Ax=x

4. Let A be a 3×3 matrix such that the following holds

$$A\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$A\mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$A\mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

(a) If the three solutions $\mathbf{x}_1, \mathbf{x}_2$, and $\mathbf{x}_3$ are the columns of a matrix X, what is AX?

(b) Suppose $\mathbf{x}_1 = (1,1,1)^T,\ \mathbf{x}_2 = (0,1,1)^T,\ \mathbf{x}_3 = (0,0,1)^T$.

i. Solve the equation $A\mathbf{x} = \mathbf{b}$ for $\mathbf{x}$ when $\mathbf{b} = (3,5,8)^T$.

ii. Find A by using the result in part (a).

5. Let A and B be two $m \times m$ invertible matrices. It is given that $A + B$ is invertible. Prove that

$$A^{-1} = (A+B)^{-1} + (A + AB^{-1}A)^{-1}.$$

6.

B = IB

7.

法一：数学归纳法

法二：A=lamda I + B (0 1 0 ; 0 0 1 ; 0 0 0)   B^3 以上的忽略（都为0）

Group 1 work

1. Suppose $T : \mathbb{Z}^3 \to \mathbb{Z}^3$ is a linear transformation, where $\mathbb{Z}$ is the set of integers. You are given that $T(\mathbf{x}) = \mathbf{y},\ T(\mathbf{y}) = \mathbf{z}$, and $T(\mathbf{z}) = \mathbf{x} + \mathbf{y}$ for certain $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{Z}^3$, where $\mathbf{x} \neq \mathbf{0}$. Prove that $\mathbf{x}, \mathbf{y}$, and $\mathbf{z}$ are linearly independent. You can use Rational root theorem without proof.

Theorem (Rational root theorem). *If a polynomial $p(x) = a_n x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_1 x + a_0$ has rational roots, those roots will be limited to $\pm\frac{p}{q}$ where p is a factor of a_0 and q is a factor of a_n .*

Solution

i) Given that $\mathbf{x} \neq \mathbf{0}$. if $\mathbf{y} = \mathbf{0}$, then

$$\mathbf{z} = T(\mathbf{y}) = T(\mathbf{0}) = \mathbf{0}$$
$$\mathbf{x} = T(\mathbf{z}) - \mathbf{y} = T(\mathbf{0}) - \mathbf{y} = \mathbf{0}\,,$$

which leads to contradiction.

if $\mathbf{z} = \mathbf{0}$, then

$$T(\mathbf{z}) = \mathbf{x} + \mathbf{y} = \mathbf{0} \Rightarrow \mathbf{y} = -\mathbf{x}$$
$$T(\mathbf{y}) = -T(\mathbf{x}) = -\mathbf{y} = \mathbf{x} = \mathbf{0}$$

which also leads to contradiction.

Hence, $\mathbf{x}, \mathbf{y}, \mathbf{z}$ are all **nonzero** vector.

ii) Suppose there exist **rational** numbers m and n so that

$\mathbf{x} = m\mathbf{y} + n\mathbf{z}$, m and n are **not all zero**.

① if $m \neq 0$ and $n = 0$, then

$$\mathbf{x} = m\mathbf{y}$$
$$T(\mathbf{x}) = mT(\mathbf{y})$$
$$\mathbf{y} = m\mathbf{z}$$
$$T(\mathbf{y}) = mT(\mathbf{z})$$
$$\mathbf{z} = m(\mathbf{x} + \mathbf{y})$$
$$\mathbf{z} = m(m+1)\mathbf{y}$$
$$\mathbf{z} = m^2(m+1)\mathbf{z}$$
$$(m^3 + m^2 - 1)\mathbf{z} = \mathbf{0}$$

By i) , $\mathbf{z} \neq \mathbf{0}$. And ± 1 are not solutions of $m^3 + m^2 - 1 = 0$.

Then, by rational root theorem,

$$m^3 + m^2 - 1 = 0$$

has no rational roots, which leads to contradiction.

② if $m = 0$ and $n \neq 0$, then

$$\mathbf{x} = n\mathbf{z}$$
$$T(\mathbf{x}) = nT(\mathbf{z})$$
$$\mathbf{y} = n(\mathbf{x} + \mathbf{y}) \quad n \neq 0 \text{ and } n \neq 1$$
$$\mathbf{x} = (\tfrac{1}{n} - 1)\mathbf{y}$$

which is similar to ① and leads to contradiction.

③ if $m \neq 0$ and $n \neq 0$, then

$$\mathbf{x} = m\mathbf{y} + n\mathbf{z}$$
$$T(\mathbf{x}) = mT(\mathbf{y}) + nT(\mathbf{z})$$
$$\mathbf{y} = m\mathbf{z} + n(\mathbf{x} + \mathbf{y})$$
$$\mathbf{y} = m\mathbf{z} + n(m\mathbf{y} + n\mathbf{z} + \mathbf{y})$$
$$(1 - mn - n)\mathbf{y} = (n^2 + m)\mathbf{z}$$

a) if $1 - mn - n \neq 0$ and $n^2 + m \neq 0$, then

$$\mathbf{y} = \tfrac{n^2+m}{1-mn-n}\mathbf{z} \Rightarrow \text{similar to ① , contradiction}$$

b) if $1 - mn - n = 0$ and $n^2 + m \neq 0$, then

$$\mathbf{z} = \mathbf{0} \Rightarrow \text{contradiction}$$

c) if $1 - mn - n \neq 0$ and $n^2 + m = 0$, then

$$\mathbf{y} = \mathbf{0} \Rightarrow \text{contradiction}$$

d) if $1 - mn - n = 0$ and $n^2 + m = 0$, then

$$\begin{cases} 1 - mn - n &= 0 \\ n^2 + m &= 0 \end{cases} \Rightarrow \begin{cases} 1 - mn - n &= 0 \\ m &= -n^2 \end{cases} \Rightarrow n^3 - n + 1 = 0$$

± 1 are not solutions of $n^3 - n + 1 = 0$.

Then, by rational root theorem,

$$n^3 - n + 1 = 0$$

has no rational roots, which leads to contradiction.

There is a similar process to suppose

$$\mathbf{y} = m\mathbf{x} + n\mathbf{z} \quad \text{or} \quad \mathbf{z} = m\mathbf{x} + n\mathbf{y}$$

All possibilities that $\mathbf{x}, \mathbf{y}$, and $\mathbf{z}$ are linearly dependent **do not exist**. Hence, $\mathbf{x}, \mathbf{y}$, and $\mathbf{z}$ are linearly independent.

2. Can you find a linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^2$ so that we have $T((1,0,3)^T) = (1,1)^T$ and $T((-2,0,-6)^T) = (2,1)^T$? Give an example if you think so or prove the impossibility otherwise.

Solution

Suppose their exists a transformation $T : \mathbb{R}^3 \to \mathbb{R}^2$ so that

$$T((1,0,3)^T) = (1,1)^T \, .$$

If the transformation is linear, then

$$T((-2,0,-6)^T) = T((-2)(1,0,3)^T) = (-2)T((1,0,3)^T) = (-2)(1,1)^T = (-2,-2)^T \neq (2,1)^T$$

So it's **impossible**.

3. (Lay 1.10, Question 9) In a certain region, about 7% of a city's population moves to the surrounding suburbs each year, and about 5% of the suburban population moves into the city. In 2015, there were 800,000 residents in the city and 500,000 in the suburbs. Set up a difference equation that describes this situation, where $\mathbf{x}_0$ is the initial population in 2015. Then estimate the populations in the city and in the suburbs two years later, in 2017. (Ignore other factors that might influence the population sizes.)

Solution

Suppose the initial (2015) city's population is x_{0c} , the initial suburban population is x_{0s} . We have

$$\mathbf{x}_0 = \begin{bmatrix} x_{0c} \\ x_{0s} \end{bmatrix}$$

Consider the next year. Suppose $\mathbf{x}_1$ is the population in 2016, city's population is x_{1c} and suburban population is x_{1s} .

According to the description,

$$x_{1c} = 93\% \ \cdot x_{0c} \ + \ \ 5\% \ \cdot x_{0s}$$

$$x_{1s} = \ 7\% \ \cdot x_{0c} \ + \ 95\% \ \cdot x_{0s}$$

Hence

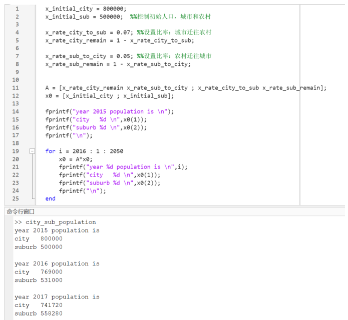
$$\mathbf{x}_1 = \begin{bmatrix} x_{1c} \\ x_{1s} \end{bmatrix} = \begin{bmatrix} 93\% \\ 7\% \end{bmatrix} x_{0c} + \begin{bmatrix} 5\% \\ 95\% \end{bmatrix} x_{0s} = \begin{bmatrix} 93\% & 5\% \\ 7\% & 95\% \end{bmatrix} \begin{bmatrix} x_{0c} \\ x_{0s} \end{bmatrix} = A\mathbf{x}_0$$

where $A = \begin{bmatrix} 93\% & 5\% \\ 7\% & 95\% \end{bmatrix}$.

Similarly for subsequent years. In general,

$$\mathbf{x}_{k+1} = A\mathbf{x}_k \qquad \text{for } k = 0, 1, 2, \cdots$$

According to the matlab running results, the city population in 2017 will be 741,720 and the suburb population will be 558,280.



```
x_initial_city = 800000;
x_initial_sub = 500000;  %%控制初始人口，城市和农村

x_rate_city_to_sub = 0.07; %%设置比率：城市迁往农村
x_rate_city_remain = 1 - x_rate_city_to_sub;

x_rate_sub_to_city = 0.05; %%设置比率：农村迁往城市
x_rate_sub_remain = 1 - x_rate_sub_to_city;


A = [x_rate_city_remain x_rate_sub_to_city ; x_rate_city_to_sub x_rate_sub_remain];
x0 = [x_initial_city ; x_initial_sub];

fprintf("year 2015 population is \n");
fprintf("city   %d \n",x0(1));
fprintf("suburb %d \n",x0(2));
fprintf("\n");

for i = 2016 : 1 : 2050
    x0 = A*x0;
    fprintf("year %d population is \n",i);
    fprintf("city   %d \n",x0(1));
    fprintf("suburb %d \n",x0(2));
    fprintf("\n");
end
```

命令行窗口

```
>> city_sub_population
year 2015 population is
city   800000
suburb 500000

year 2016 population is
city   769000
suburb 531000

year 2017 population is
city   741720
suburb 558280
```

4. Let A be a 3×3 matrix such that the following holds

$$A\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$A\mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$A\mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

(a) If the three solutions $\mathbf{x}_1, \mathbf{x}_2$, and $\mathbf{x}_3$ are the columns of a matrix X, what is AX?

(b) Suppose $\mathbf{x}_1 = (1,1,1)^T, \ \mathbf{x}_2 = (0,1,1)^T, \ \mathbf{x}_3 = (0,0,1)^T$.

i. Solve the equation $A\mathbf{x} = \mathbf{b}$ for $\mathbf{x}$ when $\mathbf{b} = (3,5,8)^T$.

ii. Find A by using the result in part (a).

Solution

(a) $AX = A \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 \end{bmatrix} = \begin{bmatrix} A\mathbf{x}_1 & A\mathbf{x}_2 & A\mathbf{x}_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3$

(b)

i. When $\mathbf{x} = 3\mathbf{x}_1 + 5\mathbf{x}_2 + 8\mathbf{x}_3$,

$$A\mathbf{x} = A(3\mathbf{x}_1 + 5\mathbf{x}_2 + 8\mathbf{x}_3) = 3A\mathbf{x}_1 + 5A\mathbf{x}_2 + 8A\mathbf{x}_3 = (3,5,8)^T = \mathbf{b}$$

Hence , there is **at least** one solution

$$\mathbf{x} = 3\mathbf{x}_1 + 5\mathbf{x}_2 + 8\mathbf{x}_3 .$$

By (a) , there exists $X = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 \end{bmatrix}$ so that $AX = I_3 \Rightarrow A$ is invertible.

By the invertible matrix theorem (Theorem 8), the equation

$$A\mathbf{x} = \mathbf{b}$$

has a **unique** solution for each $\mathbf{b}$ in $\mathbb{R}^3$.

Hence , there is **only** one solution

$$\mathbf{x} = 3\mathbf{x}_1 + 5\mathbf{x}_2 + 8\mathbf{x}_3 = \begin{bmatrix} 3 \\ 8 \\ 16 \end{bmatrix}.$$

i. #Solution 2

By Theorem 5, if A is an invertible matrix , then the equation

$$A\mathbf{x} = \mathbf{b}$$

has the **unique** solution

$$\mathbf{x} = A^{-1}\mathbf{b}$$

for each $\mathbf{b}$ in $\mathbb{R}^3$.

Hence , there is **only** one solution

$$\mathbf{x} = X\mathbf{b} = 3\mathbf{x}_1 + 5\mathbf{x}_2 + 8\mathbf{x}_3 = \begin{bmatrix} 3 \\ 8 \\ 16 \end{bmatrix}.$$

ii. In part (a) , $AX = I_3 \Rightarrow A = X^{-1}$

$$[\,X \quad I\,] = \begin{bmatrix} 1 & 0 & 0 & \vdots & 1 & 0 & 0 \\ 1 & 1 & 0 & \vdots & 0 & 1 & 0 \\ 1 & 1 & 1 & \vdots & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & \vdots & 1 & 0 & 0 \\ 0 & 1 & 0 & \vdots & -1 & 1 & 0 \\ 1 & 1 & 1 & \vdots & 0 & 0 & 1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & \vdots & 1 & 0 & 0 \\ 0 & 1 & 0 & \vdots & -1 & 1 & 0 \\ 0 & 1 & 1 & \vdots & -1 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & \vdots & 1 & 0 & 0 \\ 0 & 1 & 0 & \vdots & -1 & 1 & 0 \\ 0 & 0 & 1 & \vdots & 0 & -1 & 1 \end{bmatrix} = [\,I \quad X^{-1}\,]$$

Hence $A = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$.

5. Let A and B be two $m \times m$ invertible matrices. It is given that $A + B$ is invertible. Prove that

$$A^{-1} = (A+B)^{-1} + (A + AB^{-1}A)^{-1}.$$

Solution

$$\begin{aligned} A^{-1} &= (A+B)^{-1} \quad + (A + AB^{-1}A)^{-1} \\ \Leftrightarrow A^{-1}A &= (A+B)^{-1}A + (AI + AB^{-1}A)^{-1}A \\ \Leftrightarrow \quad I &= (A+B)^{-1}A + (A(I + B^{-1}A))^{-1}A \\ \Leftrightarrow \quad I &= (A+B)^{-1}A + (I + B^{-1}A)^{-1}A^{-1}A \\ \Leftrightarrow \quad I &= (A+B)^{-1}A + (I + B^{-1}A)^{-1}I \\ \Leftrightarrow \quad I &= (A+B)^{-1}A + (I + B^{-1}A)^{-1}B^{-1}B \\ \Leftrightarrow \quad I &= (A+B)^{-1}A + (B(I + B^{-1}A))^{-1}B \\ \Leftrightarrow \quad I &= (A+B)^{-1}A + (B+A)^{-1}B \\ \Leftrightarrow \quad I &= (A+B)^{-1}(A+B)\,, \end{aligned}$$

which is easy to prove by definition.

exercise 3

Group 1 : 3 4 6 7 9

1. Let A, X be $n \times n$ matrices and $AX = I$, where I is the $n \times n$ identify matrix. Is A invertible? If true, calculate A^{-1} ; otherwise, give a counterexample.

   定义 $AX = XA = I \Rightarrow A$ is invertible

   本题不满足定义

   for any $\mathbf{b}$ in $\mathbb{R}^n$, if $\mathbf{b} = k_1 A_1 + \cdots + k_n A_n = A\mathbf{k}$, then span(A) =$\mathbb{R}^n$

   $\mathbf{b} = I\mathbf{b} = (AX)\mathbf{b} = A(X\mathbf{b})$

   hence $\mathbf{k} = X\mathbf{b}$

   is what we want

2. Let X be a $m \times m$ matrix. It is given that:

   (i) xii ≥ 1 for all i, where xij is the (i, j)-entry of matrix X, and

   (ii) P i=j x 2 ij < 1. Show that X is invertible. You can use the following results without proof.

   Theorem. Let $u, v \in Rm$. If $uTv = 0$ and v is a non-zero vector, then u must be a non-zero vector.

   Theorem (Cauchy–Schwarz inequality). For any $u, v \in Rn$, $(Pni = 1uivi)2 \leq Pni = 1u2iPni = 1v2i$.

   (Hint: You need to show that $X\mathbf{y} \neq \mathbf{0}$ for any non-zero vector $y \in Rm$. In that case, the equation $Xy = 0$ has only the trivial solution and hence X is invertible.)

   要证 for any $\mathbf{y} \neq \mathbf{0}$, 有 $X\mathbf{y} \neq \mathbf{0}$ (只有 0 解)

   只需证 $(X\mathbf{y})^T\mathbf{y} \neq \mathbf{0}$

   和 $\mathbf{y} \neq \mathbf{0}$ （已知）

   $(X\mathbf{y})^T\mathbf{y} = \sum_{i,j=1}^{n}$

   这题最后一步可以反证施瓦兹定理

3. Let Y be a matrix of size $n \times n$ such that $\sum_{j=1}^{n} |y_{ij}| < 1$ for all i, where y_{ij} is the (i, j)-th entry of matrix Y. Show that $I - Y$ is invertible.

   $$(Hint: \ Similar\ to\ Question\ 2,\ you\ need\ to\ show\ that\ the\ zero\ vector\ \mathbf{0}\ is\ the\ only\ solution\ to\ the\ equation\ (I - Y)\mathbf{x} = \mathbf{0}.)$$

4. Let X be a matrix of size $2n \times 2n$ which has the following form

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

where A, B, C, and D are $n \times n$ matrices which commute with each other, i.e., $AB = BA$, $AC = CA$, $AD = DA$, $BC = CB$, $BD = DB$, and $CD = DC$. Show that $AD - BC$ is invertible if and only if X^{-1} exists. You may use the following result without proof.

Theorem. Let A and B be two $m \times m$ matrices. A and B are invertible if and only if AB is invertible.

5. (Lay 2.4, Questions 15 and 18)

   (a) Suppose A11 is invertible. Find X and Y such that A11 A12 A21 A22 = I 0 X I A11 0 0 S I Y 0 I where S = A22 − A21A −1 11 A12. The matrix S is called the Schur complement of A11. Likewise, if A22 is invertible, the matrix A11 − A12A −1 22 A21 is called the Schur complement of A22.

   (b) Let X be an m × n data matrix such that XT X is invertible, and let M = Im − X(XT X) −1XT . Add a column x0 to the data and form W = X x0 . Compute WT W and show that the Schur complement of XT X can be written in the form x T 0 Mx0.

   此处 X^T和X 不可逆，因此相乘不能分开

6. (Lay 2.5, Question 24) Suppose $A = QR$, where Q and R are $n \times n$, R is invertible and upper triangular, and Q has the property that $Q^TQ = I$. Show that for each $\mathbf{b}$ in $\mathbb{R}^n$, the equation $A\mathbf{x} = \mathbf{b}$ has a unique solution. What computations with Q and R will produce the solution?

7. (Lay 2.5, Question 25) Suppose $A = UDV^T$, where U and V are $n \times n$ matrices with the property that $U^TU = I$ and $V^TV = I$, and where D is a diagonal matrix with positive numbers $\sigma_1, \sigma_2, \ldots, \sigma_n$ on the diagonal. Show that A is invertible, and find a formula of A^{-1} .

8. Use partitioned matrices to prove the two parts by induction.

   (a) The product of two lower triangular matrices is also lower triangula

   (b) The n × n (n ≥ 2) matrix A shown below is invertible and B is its inverse. A = □ □ □ □ □ □ □ 1 0 0 · · · 0 1 1 0 · · · 0 1 1 1 · · · 0 . . . . 1 1 1 · · · 1 □ □ □ □ □ □ □ , B = □ □ □ □ □ □ □ 1 0 0 · · · 0 −1 1 0 · · · 0 0 −1 1 · · · 0 . . . . 0 0 0 · · · 1 □ □ □ □ □ □ □ . You may find the following information helpful when answering Questions 9. No proof is required when using the facts below. Fact. The maximum number of linearly independent rows in a matrix A is called the row rank of A, and the maximum number of linearly independent columns in A is called the column rank of A. For any matrix A, we must have the row rank of A equal to the column rank of A. The rank of matrix A is defined as follows: Rank of A = Row rank of A = Column rank of A. Only a zero matrix has rank zero. Moreover, we also have the following. Theorem. Let B be a n × n matrix. The equation Bx = 0 has non-trivial solutions if and only if the rank of B is less than n.

9. Let A be a $n \times n$ matrix such that:

   (i) all of its entries, a_{ij} , are either 0 or 1, and

   (ii) $a_{ii} = 0$ and $a_{ij} + a_{ji} = 1$ for all $1 \le i < j \le n$.

   Prove that the rank of matrix A must be at least $n - 1$.

Group 1 work

3. Let Y be a matrix of size $n \times n$ such that $\sum_{j=1}^{n} |y_{ij}| < 1$ for all i , where y_{ij} is the (i, j)-th entry of matrix Y . Show that $I - Y$ is invertible.

   $(Hint: \ Similar \ to \ Question \ 2, \ you \ need \ to \ show \ that \ the \ zero \ vector \ \mathbf{0} \ is \ the \ only \ solution \ to \ the \ equation \ (I - Y)\mathbf{x} = \mathbf{0}.)$

Solution:

For any $\mathbf{x} \neq \mathbf{0}$,

$$(I - Y)\mathbf{x} = \begin{bmatrix} 1 - y_{11} & -y_{12} & \cdots & -y_{1n} \\ -y_{21} & 1 - y_{22} & \cdots & -y_{2n} \\ \vdots & & & \vdots \\ -y_{n1} & -y_{n2} & \cdots & 1 - y_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_1 - \sum_{j=1}^{n} x_j y_{1j} \\ x_2 - \sum_{j=1}^{n} x_j y_{2j} \\ \vdots \\ x_n - \sum_{j=1}^{n} x_j y_{nj} \end{bmatrix}.$$

Suppose for any integer $i \in [1, n]$, $|x_i| \le |x_{max}|$ $(max \in \mathbb{Z} \ and \ max \in [1, n])$.

$$\mathbf{x} \neq \mathbf{0} \Rightarrow |x_{max}| > 0 \Rightarrow -|x_{max}| < 0 \Rightarrow -|x_{max}| \sum_{j=1}^{n} |y_{maxj}| > -|x_{max}|$$

$$
\begin{aligned}
|x_{max} - \sum_{j=1}^{n} x_j y_{maxj}| &\geq |x_{max}| - |\sum_{j=1}^{n} x_j y_{maxj}| \\
&\geq |x_{max}| - \sum_{j=1}^{n} |x_j||y_{maxj}| \\
&\geq |x_{max}| - |x_{max}| \sum_{j=1}^{n} |y_{maxj}| \\
&> |x_{max}| - |x_{max}| = 0
\end{aligned}
$$

Then for any $\mathbf{x} \neq \mathbf{0}$, $(I-Y)\mathbf{x} \neq \mathbf{0} \Rightarrow I-Y$ is invertible.

要证 $I-Y$ is invertible

只需证

for any $\mathbf{x} \neq \mathbf{0} \Rightarrow (I-Y)\mathbf{x} \neq \mathbf{0}$

放缩过程相似（选 x 最大项所在行讨论）

列昂惕夫模型变式（列昂惕夫是 column sum < 1，这题是 row sum < 1）

solve $(I-Y)\mathbf{x} = \mathbf{0}$.

$(I-Y)\mathbf{x} = \mathbf{0}$

$(I + Y^2 + Y^3 + \cdots + Y^m)(I-Y)\mathbf{x} = \mathbf{0}$

$(I - Y^{m+1})\mathbf{x} = \mathbf{0}$ m 任意正整数

when $m \to \infty$, $Y^{m+1} \to 0$, $\mathbf{x} = \mathbf{0}$

证明：when $m \to \infty$, $Y^{m+1} \to 0$

$$
Y = \begin{bmatrix} y_{11} & \cdots & y_{1n} \\ \vdots & & \vdots \\ y_{i1} & & y_{in} \\ \vdots & & \vdots \\ y_{n1} & \cdots & y_{nn} \end{bmatrix}, \quad \sum_{j=1}^{n} |y_{ij}| < 1 \text{ for all } i
$$

$|y_{11}| + |y_{12}| + \cdots + |y_{1n}| < 1$

$|y_{21}| + |y_{22}| + \cdots + |y_{2n}| < 1$

$\vdots$

$|y_{n1}| + |y_{n2}| + \cdots + |y_{nn}| < 1$

并且隐藏条件（imply）：每个元素（every entry）绝对值不大于 1

$$Y^2 = \begin{bmatrix} y_{11} & \cdots & \cdots & y_{1n} \\ \vdots & & & \vdots \\ y_{i1} & y_{i2} & \cdots & y_{in} \\ \vdots & & & \vdots \\ y_{n1} & \cdots & \cdots & y_{nn} \end{bmatrix} \begin{bmatrix} y_{11} & \cdots & y_{1n} \\ \vdots & & \vdots \\ y_{i1} & & y_{in} \\ \vdots & & \vdots \\ y_{n1} & \cdots & y_{nn} \end{bmatrix}$$

$$= \begin{bmatrix} Y_{11}^2 & \cdots & \cdots & Y_{1n}^2 \\ \vdots & & & \vdots \\ y_{i1}y_{11} + y_{i2}y_{21} + \cdots + y_{in}y_{n1} & y_{i1}y_{12} + y_{i2}y_{22} + \cdots + y_{in}y_{n2} & \cdots & y_{i1}y_{1n} + y_{i2}y_{2n} + \cdots + y_{in}y_{nn} \\ \vdots & & & \vdots \\ Y_{n1}^2 & \cdots & \cdots & Y_{nn}^2 \end{bmatrix}$$

记 A 第 i 行 的绝对值之和为 $absRsum(A_i)$

$absRsum(Y_i^2) = |y_{i1}y_{11} + y_{i2}y_{21} + \cdots + y_{in}y_{n1}| + |y_{i1}y_{12} + y_{i2}y_{22} + \cdots + y_{in}y_{n2}| + \cdots + |y_{i1}y_{1n} + y_{i2}y_{2n} + \cdots + y_{in}y_{nn}|$

$\leq |y_{i1}|(|y_{11}| + |y_{12}| + \cdots + |y_{1n}|) + |y_{i2}|(|y_{21}| + |y_{22}| + \cdots + |y_{2n}|) + \cdots + |y_{in}|(|y_{n1}| + |y_{n2}| + \cdots + |y_{nn}|)$

$< |y_{i1}| + |y_{i2}| + \cdots + |y_{in}| = absRsum(Y_i)$

极限是 0

为什么是0?

可以设 Y 的最大 row absolute sum 为 $max(max < 1)$

$absRsum(Y_i^2) \leq max \cdot absRsum(Y_i)$

那么 Y^2 的最大可能的 row sum 为 $(max)^2$

那么 Y^{m+1} 的最大可能的 row sum 为 $(max)^{m+1}$

$\lim_{m \to \infty} (max)^{m+1} = 0$

于是 m 趋于无穷时Y的所有行都趋于 0， 即 Y 趋于 0 矩阵

法二：

要证 $I - Y$ is invertible

只需证

for any $\mathbf{x} \neq \mathbf{0} \Rightarrow (I - Y)\mathbf{x} \neq \mathbf{0}$

放缩过程相似（选 x 最大项所在行讨论）

4. Let X be a matrix of size $2n \times 2n$ which has the following form

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

where A, B, C, and D are $n \times n$ matrices which commute with each other, i.e., $AB = BA$, $AC = CA$, $AD = DA$, $BC = CB$, $BD = DB$, and $CD = DC$. Show that $AD - BC$ is invertible if and only if X^{-1} exists. You may use the following result without proof.

Theorem. Let A and B be two $m \times m$ matrices. A and B are invertible if and only if AB is invertible.

Solution:

① Y is invertible $\Rightarrow$ X is invertible:

$$X\begin{bmatrix} D & -B \\ -C & A \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}\begin{bmatrix} D & -B \\ -C & A \end{bmatrix} = \begin{bmatrix} AD-BC & 0 \\ 0 & AD-BC \end{bmatrix} = \begin{bmatrix} Y & 0 \\ 0 & Y \end{bmatrix}$$

Hence Y is invertible $\Rightarrow$ X is invertible.

② X is invertible $\Rightarrow$ Y is invertible:

Suppose X is invertible , $Y\omega = \mathbf{0}$, then

$X\begin{bmatrix} D\omega \\ -C\omega \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}\begin{bmatrix} D\omega \\ -C\omega \end{bmatrix} = \begin{bmatrix} Y\omega \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix} \Rightarrow C\omega = \mathbf{0}$, $D\omega = \mathbf{0}$

$X\begin{bmatrix} -B\omega \\ A\omega \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}\begin{bmatrix} -B\omega \\ A\omega \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ Y\omega \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix} \Rightarrow A\omega = \mathbf{0}$, $B\omega = \mathbf{0}$

$X\begin{bmatrix} \omega \\ \omega \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}\begin{bmatrix} \omega \\ \omega \end{bmatrix} = \begin{bmatrix} A\omega + B\omega \\ C\omega + D\omega \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix} \Rightarrow \omega = \mathbf{0} \Rightarrow Y$ is invertible.

Hence X is invertible $\Rightarrow$ Y is invertible

①② $\Rightarrow$ $AD - BC$ is invertible if and only if X^{-1} exists.

灵感来源：2.2 Theorem 4

Theorem 4 二阶方阵可逆性判别 & 逆矩阵求法

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If $ad - bc \neq 0$, then A is invertible and

$A^{-1} = \dfrac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

If $ad - bc = 0$, then A is not invertible.

构造

$X^{-1} = (AD - BC)^{-1}\begin{bmatrix} D & -B \\ -C & A \end{bmatrix}$

这个构造是 为了 联系 AD-BC 构造出来的

②的证明用了 AB invertible 等价于 A B 分别 invertible

验证：经过一波计算发现 $X^{-1}X = I_{2n}$

$X^{-1}\ exists \ \Leftrightarrow\ (AD-BC)^{-1}\ exists \ \Leftrightarrow\ AD - BC$ is invertible （不严谨）

顺带一提：本题给ABCD commute with each other，因为之前只定义了矩阵相乘（把右边矩阵的列中元素作为权对左边矩阵的列进行组合，生成新列），没有定义块矩阵相乘。验证X^{-1}构造正确时，会出现左右乘的顺序问题，给这个条件就避免了这个问题（随便怎么乘都一样）。

6. (Lay 2.5, Question 24) Suppose $A = QR$, where Q and R are $n \times n$, R is invertible and upper triangular, and Q has the property that $Q^TQ = I$. Show that for each $\mathbf{b}$ in $\mathbb{R}^n$, the equation $A\mathbf{x} = \mathbf{b}$ has a unique solution. What computations with Q and R will produce the solution?

Solution:

$Q^TQ = I \Rightarrow Q$ is invertible and $Q^{-1} = Q^T$

R, Q are invertible $\Rightarrow A$ is invertible $\Rightarrow$ the equation $A\mathbf{x} = \mathbf{b}$ has a unique solution $A^{-1}\mathbf{b} = R^{-1}Q^T\mathbf{b}$

A good algorithm for finding $\mathbf{x}$ is to compute $Q^T\mathbf{b}$ and then row reduce the matrix $[\, R \quad Q^T\mathbf{b} \,]$. The reduction is fast in this case because R is a triangular matrix.

$R\mathbf{x} = Q^T\mathbf{b}$ 好解

7. (Lay 2.5, Question 25) Suppose $A = UDV^T$, where U and V are $n \times n$ matrices with the property that $U^TU = I$ and $V^TV = I$, and where D is a diagonal matrix with positive numbers $\sigma_1, \sigma_2, \ldots, \sigma_n$ on the diagonal. Show that A is invertible, and find a formula of A^{-1}.

Solution:

$U^TU = I$ and $V^TV = I \Rightarrow U$ is invertible and V is invertible.

D is a diagonal matrix with positive numbers $\sigma_1, \sigma_2, \ldots, \sigma_n$ on the diagonal $\Rightarrow D$ is invertible.

Hence A is invertible and $A^{-1} = VD^{-1}U^T$, where D^{-1} is a diagonal matrix with positive numbers $\sigma_1^{-1}, \sigma_2^{-1}, \ldots, \sigma_n^{-1}$ on the diagonal.

You may find the following information helpful when answering Questions 9. No proof is required when using the facts below. Fact. The maximum number of linearly independent rows in a matrix A is called the row rank of A, and the maximum number of linearly independent columns in A is called the column rank of A. For any matrix A, we must have the row rank of A equal to the column rank of A. The rank of matrix A is defined as follows: Rank of A = Row rank of A = Column rank of A. Only a zero matrix has rank zero. Moreover, we also have the following. Theorem. Let B be a n × n matrix. The equation Bx = 0 has non-trivial solutions if and only if the rank of B is less than n.

9. Let A be a $n \times n$ matrix such that:

   (i) all of its entries, a_{ij} , are either 0 or 1, and

   (ii) $a_{ii} = 0$ and $a_{ij} + a_{ji} = 1$ for all $1 \le i < j \le n$.

   Prove that the rank of matrix A must be at least $n - 1$.

Solution:

$$A + A^T + I = J = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \cdots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}$$

Assume $rank(A) \le n-2 \Rightarrow rank(\begin{bmatrix} A \\ J \end{bmatrix}) \le n-1$, then $\exists\, \mathbf{x} \ne \mathbf{0}$ such that $\begin{bmatrix} A \\ J \end{bmatrix}\mathbf{x} = \mathbf{0}$.

$\begin{bmatrix} A \\ J \end{bmatrix}\mathbf{x} = \mathbf{0} \Leftrightarrow \begin{cases} A\mathbf{x} = \mathbf{0} \\ J\mathbf{x} = \mathbf{0} \end{cases}$

When $\mathbf{x} \ne \mathbf{0}$ and $\begin{cases} A\mathbf{x} = \mathbf{0} \\ J\mathbf{x} = \mathbf{0} \end{cases}$,

$$\begin{aligned} 0 < x_1^2 + x_2^2 + \cdots + x_n^2 &= \mathbf{x}^T\mathbf{x} \\ &= \mathbf{x}^T I \mathbf{x} \\ &= \mathbf{x}^T(J - A - A^T)\mathbf{x} \\ &= \mathbf{x}^T J\mathbf{x} - \mathbf{x}^T A\mathbf{x} - \mathbf{x}^T A^T \mathbf{x} \\ &= \mathbf{x}^T J\mathbf{x} - \mathbf{x}^T A\mathbf{x} - (A\mathbf{x})^T \mathbf{x} \\ &= 0 - 0 - 0 \\ &= 0 \end{aligned}$$

which leads to contradiction.

Hence, there **DO NOT** exist $\mathbf{x} \ne \mathbf{0}$ such that $\begin{bmatrix} A \\ J \end{bmatrix}\mathbf{x} = \mathbf{0} \Rightarrow rank(A) \ge n-1$.

思路：可以发现最多只有 1 个 全零 行或列

无全零行 - row reduced echelon form 有 n 个 pivot column - n秩

有一个全零行 - row reduced echelon form 有 n-1 个 pivot column - n-1秩

$A + A^T + I = J$ j 是所有元素为1的矩阵

$rank\ A + dim\ Nul\ A = n$

exercise 4

Group 1 : 1,3,5,6,7

1. Calculate the following determinants. They are all $n \times n$ matrices if no additional explanation.

   a. Determinant of an anti-symmetric matrix, where n is odd.

$$\begin{vmatrix} 0 & a_{12} & \cdots & a_{1n} \\ -a_{12} & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{1n} & -a_{2n} & \cdots & 0 \end{vmatrix}$$

   b. $$\begin{vmatrix} x & -1 & 0 & \cdots & 0 & 0 \\ 0 & x & -1 & \cdots & 0 & 0 \\ 0 & 0 & x & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & x & -1 \\ a_n & a_{n-1} & a_{n-2} & \cdots & a_2 & x+a_1 \end{vmatrix}$$

   c. $$\begin{vmatrix} 1 & 3 & 3 & \cdots & 3 \\ 3 & 2 & 3 & \cdots & 3 \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ 3 & 3 & \cdots & n-1 & 3 \\ 3 & 3 & \cdots & 3 & n \end{vmatrix}$$

2. Calculate the following determinants. They are all n × n matrices if no additional explanation.

   a. 1 + x 2 · · · n 1 2 + x · · · n . . . . 1 2 · · · n + x

   b. a1 − b1 a1 − b2 · · · a1 − bn a2 − b1 a2 − b2 · · · a2 − bn . . . . an − b1 an − b2 · · · an − bn

   (Hint: You may rewrite the original matrix as the product of two matrices.)

3. Let A, B, C, D and I be $n \times n$ matrices, and O is an $n \times n$ zero matrix. Use the definition and properties of a determinant to justify the following formulas.

   a. $\det \begin{bmatrix} A & O \\ O & I \end{bmatrix} = \det A$

   b. $\det \begin{bmatrix} I & O \\ C & D \end{bmatrix} = \det D$

   c. $\det \begin{bmatrix} A & O \\ C & D \end{bmatrix} = \det \begin{bmatrix} A & B \\ O & D \end{bmatrix} = (\det A)(\det D)$

4. For any real square matrix A, the adjugate matrix, adj A, has as its elements the cofactors of the transpose of A, i.e., A(adj A) = (adj A)A = (det A)I.

   a. Suppose that A is an invertible square matrix. Show that (adj (A T))−1 = (adj (A −1))T .

   b. Suppose that adj (AT) is invertible. Prove that A is invertible.

5. a. For an $n \times n$ matrix $A = [\, aij \,]$, suppose its k-th row can be written as the summation of two arrays, i.e., $a_{kj} = \theta_{kj} + \eta_{kj}$ for $j = 1, \cdots, n$. Denote A_θ as the matrix which replace the k-th row in A by the array θ_{kj} for $j = 1, \cdots, n$, and remain other elements unchanged. Denote A_η similarly. Prove that $detA = detA_\theta + detA_\eta$.

b. Use the result in (a), calculate the following deternimant: $\begin{vmatrix} x & b & b & \cdots & b \\ c & x & b & \cdots & b \\ c & c & x & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c & c & c & \cdots & x \end{vmatrix}$.

6. An anti-symmetric matrix is a square matrix whose transpose equals its negative. For an $n \times n$ antisymmetric matrix $A = [\, a_{ij} \,]$, compute det A when n is even and $a_{ij} = 1$ for $1 \le i < j \le n$.

7. a. Calculate the determinant: $\begin{vmatrix} 1+x_1^2 & x_1x_2 & \cdots & x_1x_n \\ x_2x_1 & 1+x_2^2 & \cdots & x_2x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_nx_1 & x_nx_2 & \cdots & 1+x_n^2 \end{vmatrix}$.

(Hint: You may add a row and a column to the matrix, such that the adding terms do not change the value of the determinant.)

b. Prove: $\begin{vmatrix} a_{11}+x_1 & \cdots & a_{1n}+x_n \\ \vdots & \vdots & \vdots \\ a_n+x_1 & \cdots & a_{nn}+x_n \end{vmatrix} = det\ A + \sum_{j=1}^{n} x_j(-1)^j\ det\ D_j$, with $A = [\, a_{ij} \,]$ and D_j is the matrix fromed by deleting the $(j+1)$-th column of $\begin{pmatrix} -1 & a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ -1 & a_{n1} & \cdots & a_{nn} \end{pmatrix}$.

Group 1 work

1. Calculate the following determinants. They are all $n \times n$ matrices if no additional explanation.

a. Determinant of an anti-symmetric matrix, where n is odd.

$$\begin{vmatrix} 0 & a_{12} & \cdots & a_{1n} \\ -a_{12} & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{1n} & -a_{2n} & \cdots & 0 \end{vmatrix}$$

b. $\begin{vmatrix} x & -1 & 0 & \cdots & 0 & 0 \\ 0 & x & -1 & \cdots & 0 & 0 \\ 0 & 0 & x & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & x & -1 \\ a_n & a_{n-1} & a_{n-2} & \cdots & a_2 & x+a_1 \end{vmatrix}$

c. $\begin{vmatrix} 1 & 3 & 3 & \cdots & 3 \\ 3 & 2 & 3 & \cdots & 3 \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ 3 & 3 & \cdots & n-1 & 3 \\ 3 & 3 & \cdots & 3 & n \end{vmatrix}$

solution:

a.
$det(A) = det(-A^T) = (-1)^n det(A^T) = -det(A) \Rightarrow det(A) = 0$

b.
Define $\begin{vmatrix} x & -1 & 0 & \cdots & 0 & 0 \\ 0 & x & -1 & \cdots & 0 & 0 \\ 0 & 0 & x & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & x & -1 \\ a_n & a_{n-1} & a_{n-2} & \cdots & a_2 & x+a_1 \end{vmatrix}$ as A_n, then we have

$$\begin{aligned} A_n &= xA_{n-1} + (-1)^{n+1}a_n(-1)^{n-1} \\ &= a_n + xA_{n-1} \\ &= a_n + x(a_{n-1} + xA_{n-2}) \\ &= a_n + xa_{n-1} + x^2(a_{n-2} + xA_{n-3}) \\ &= a_n + xa_{n-1} + x^2a_{n-2} + \cdots + x^{n-2}(a_2 + xA_1) \\ &= a_n + xa_{n-1} + x^2a_{n-2} + \cdots + x^{n-2}a_2 + x^{n-1}a_1 + x^n \end{aligned}$$

c.

$$\begin{vmatrix} 1 & 3 & 3 & \cdots & 3 \\ 3 & 2 & 3 & \cdots & 3 \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ 3 & 3 & \cdots & n-1 & 3 \\ 3 & 3 & \cdots & 3 & n \end{vmatrix} = \begin{vmatrix} -2 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 0 & 0 & 0 & \cdots & 0 \\ 3 & 3 & 3 & 3 & 3 & \cdots & 3 \\ 0 & 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \cdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & n-4 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & n-3 \end{vmatrix}$$

$$= \begin{vmatrix} -2 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 3 & 3 & 3 & \cdots & 3 \\ 0 & 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \cdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & n-4 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & n-3 \end{vmatrix}$$

$$= 6 \times (n-3)!$$

3. Let A, B, C, D and I be $n \times n$ matrices, and O is an $n \times n$ zero matrix. Use the definition and properties of a determinant to justify the following formulas.

a. det $\begin{bmatrix} A & O \\ O & I \end{bmatrix} =$ det A

b. det $\begin{bmatrix} I & O \\ C & D \end{bmatrix} =$ det D

c. det $\begin{bmatrix} A & O \\ C & D \end{bmatrix} =$ det $\begin{bmatrix} A & B \\ O & D \end{bmatrix} =$ (det A)(det D)

solution:

a.

$$det\begin{bmatrix} A & O \\ O & I \end{bmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} & 0 & 0 & \cdots \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 0 & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 & \cdots \\ 0 & 0 & \cdots & 0 & 1 & 0 & \cdots \\ 0 & 0 & \cdots & 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots \end{vmatrix}$$

$$= \cdots$$

$$= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} & 0 & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 \\ 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$$= det\ A.$$

b.

$$det\begin{bmatrix} I & O \\ C & D \end{bmatrix} = \begin{vmatrix} 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \\ c_{11} & c_{12} & \cdots & c_{1n} & d_{11} & d_{12} & \cdots & d_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} & d_{21} & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} & d_{n1} & d_{n2} & \cdots & d_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & d_{11} & d_{12} & \cdots & d_{1n} \\ 0 & 0 & \cdots & 0 & d_{21} & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & d_{n1} & d_{n2} & \cdots & d_{nn} \end{vmatrix}$$

$$= det\ D.$$

c.

$$det\begin{bmatrix} A & O \\ C & D \end{bmatrix} = det(\begin{bmatrix} A & O \\ O & I \end{bmatrix}\begin{bmatrix} I & O \\ C & D \end{bmatrix})$$

$$= (det\begin{bmatrix} A & O \\ O & I \end{bmatrix})(det\begin{bmatrix} I & O \\ C & D \end{bmatrix})$$

$$= (det\ A)(det\ D).$$

solution 1

$$det\begin{bmatrix} A & B \\ O & D \end{bmatrix} = det(\begin{bmatrix} I & B \\ O & D \end{bmatrix}\begin{bmatrix} A & O \\ O & I \end{bmatrix})$$

$$= (det\begin{bmatrix} I & B \\ O & D \end{bmatrix})(det\begin{bmatrix} A & O \\ O & I \end{bmatrix})$$

$$= (det\ D)(det\ A)$$

$$= (det\ A)(det\ D).$$

solution 2

$$det\begin{bmatrix} A & B \\ O & D \end{bmatrix} = det\begin{bmatrix} A & B \\ O & D \end{bmatrix}^T$$

$$= det\begin{bmatrix} A^T & O \\ B^T & D^T \end{bmatrix}$$

$$= (det\ A^T)(det\ D^T)$$

$$= (det\ A)(det\ D).$$

Hence det $\begin{bmatrix} A & O \\ C & D \end{bmatrix} =$ det $\begin{bmatrix} A & B \\ O & D \end{bmatrix} =$ (det A)(det D) .

5. a. For an $n \times n$ matrix $A = [\, aij \,]$, suppose its k-th row can be written as the summation of two arrays, i.e., $a_{kj} = \theta_{kj} + \eta_{kj}$ for $j = 1, \cdots, n$. Denote A_θ as the matrix which replace the k-th row in A by the array θ_{kj} for $j = 1, \cdots, n$, and remain other elements unchanged. Denote A_η similarly. Prove that $det\ A = det\ A_\theta + det\ A_\eta$.

solution:

$$det\ A = \sum_{j=1}^{n} a_{kj} C_{kj}$$

$$= \sum_{j=1}^{n} (\theta_{kj} + \eta_{kj}) C_{kj}$$

$$= \sum_{j=1}^{n} \theta_{kj} C_{kj} + \sum_{j=1}^{n} \eta_{kj} C_{kj}$$

$$= det\ A_\theta + det\ A_\eta$$

b. Use the result in (a), calculate the following deternimant: $\begin{vmatrix} x & b & b & \cdots & b \\ c & x & b & \cdots & b \\ c & c & x & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c & c & c & \cdots & x \end{vmatrix}$.

solution:

Define $A^n = \overbrace{\begin{vmatrix} x & b & b & \cdots & b \\ c & x & b & \cdots & b \\ c & c & x & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c & c & c & \cdots & x \end{vmatrix}}^{n}$, then we have

$$A^n = \begin{vmatrix} x-b & b & b & \cdots & b \\ c & x & b & \cdots & b \\ c & c & x & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c & c & c & \cdots & x \end{vmatrix} + \begin{vmatrix} b & b & b & \cdots & b \\ c & x & b & \cdots & b \\ c & c & x & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c & c & c & \cdots & x \end{vmatrix}$$

$$= \begin{vmatrix} x-b & 0 & 0 & \cdots & 0 \\ c & x & b & \cdots & b \\ c & c & x & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c & c & c & \cdots & x \end{vmatrix} + \begin{vmatrix} b & b & b & \cdots & b \\ c & x & b & \cdots & b \\ c & c & x & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c & c & c & \cdots & x \end{vmatrix}$$

$$= (x-b)A^{n-1} + b\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ c & x & b & \cdots & b \\ c & c & x & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c & c & c & \cdots & x \end{vmatrix}$$

$$= (x-b)A^{n-1} + b\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & x-c & b-c & \cdots & b-c \\ 0 & 0 & x-c & \cdots & b-c \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & x-c \end{vmatrix}$$

$$= b(x-c)^{n-1} + (x-b)A^{n-1}$$

$$= b(x-c)^{n-1} + (x-b)(b(x-c)^{n-2} + (x-b)A^{n-2})$$

$$= b(x-c)^{n-1} + b(x-b)(x-c)^{n-2} + (x-b)^2A^{n-2}$$

$$= b(x-c)^{n-1} + b(x-b)(x-c)^{n-2} + b(x-b)^2(x-c)^{n-3} + \cdots + b(x-b)^{n-2}(x-c) + (x-b)^{n-1}x$$

6. An anti-symmetric matrix is a square matrix whose transpose equals its negative. For an $n \times n$ antisymmetric matrix $A = [\ a_{ij}\ ]$, compute det A when n is even and $a_{ij} = 1$ for $1 \leq i < j \leq n$.

solution:

$$A^n = \begin{vmatrix} 0 & 1 & 1 & \cdots & 1 \\ -1 & 0 & 1 & \cdots & 1 \\ -1 & -1 & 0 & \cdots & 1 \\ -1 & -1 & -1 & \ddots & 1 \\ -1 & -1 & -1 & \cdots & 0 \end{vmatrix}$$

$$= \begin{vmatrix} -1 & 1 & 1 & \cdots & 1 \\ -1 & 0 & 1 & \cdots & 1 \\ -1 & -1 & 0 & \cdots & 1 \\ -1 & -1 & -1 & \ddots & 1 \\ -1 & -1 & -1 & \cdots & 0 \end{vmatrix} + \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 0 & 1 & \cdots & 1 \\ 0 & -1 & 0 & \cdots & 1 \\ 0 & -1 & -1 & \ddots & 1 \\ 0 & -1 & -1 & \cdots & 0 \end{vmatrix}$$

$$= \begin{vmatrix} -1 & 0 & 0 & \cdots & 0 \\ -1 & -1 & 0 & \cdots & 0 \\ -1 & -2 & -1 & \cdots & 0 \\ -1 & -2 & -2 & \ddots & 0 \\ -1 & -2 & -2 & \cdots & -1 \end{vmatrix} + A^{n-1}$$

$$= (-1)^n + (-1)^{n-1} + \cdots + (-1)^2 + 0$$

$$= 1$$

7. a. Calculate the determinant: $\begin{vmatrix} 1+x_1^2 & x_1x_2 & \cdots & x_1x_n \\ x_2x_1 & 1+x_2^2 & \cdots & x_2x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_nx_1 & x_nx_2 & \cdots & 1+x_n^2 \end{vmatrix}$.

(Hint: You may add a row and a column to the matrix, such that the adding terms do not change the value of the determinant.)

solution:

$$\begin{vmatrix} 1+x_1^2 & x_1x_2 & \cdots & x_1x_n \\ x_2x_1 & 1+x_2^2 & \cdots & x_2x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_nx_1 & x_nx_2 & \cdots & 1+x_n^2 \end{vmatrix} = \begin{vmatrix} 1 & x_1 & x_2 & \cdots & x_n \\ 0 & 1+x_1^2 & x_1x_2 & \cdots & x_1x_n \\ 0 & x_2x_1 & 1+x_2^2 & \cdots & x_2x_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & x_nx_1 & x_nx_2 & \cdots & 1+x_n^2 \end{vmatrix}$$

$$= \begin{vmatrix} 1 & x_1 & x_2 & \cdots & x_n \\ -x_1 & 1 & 0 & \cdots & 0 \\ -x_2 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -x_n & 0 & 0 & \cdots & 1 \end{vmatrix}$$

$$= \begin{vmatrix} 1+x_1^2+x_2^2+\cdots+x_n^2 & 0 & 0 & \cdots & 0 \\ -x_1 & 1 & 0 & \cdots & 0 \\ -x_2 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -x_n & 0 & 0 & \cdots & 1 \end{vmatrix}$$

$$= 1 + x_1^2 + x_2^2 + \cdots + x_n^2$$

b. Prove: $\begin{vmatrix} a_{11}+x_1 & \cdots & a_{1n}+x_n \\ \vdots & \vdots & \vdots \\ a_n+x_1 & \cdots & a_{nn}+x_n \end{vmatrix} = det\ A + \sum_{j=1}^{n} x_j(-1)^j\ det\ D_j$, with $A = [\ aij\ ]$ and D_j is the matrix fromed by deleting the $(j+1)$-th column of $\begin{pmatrix} -1 & a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ -1 & a_{n1} & \cdots & a_{nn} \end{pmatrix}$.

solution:

$$\begin{vmatrix} a_{11}+x_1 & \cdots & a_{1n}+x_n \\ \vdots & \vdots & \vdots \\ a_n+x_1 & \cdots & a_{nn}+x_n \end{vmatrix} = \begin{vmatrix} 1 & x_1 & \cdots & x_n \\ 0 & a_{11}+x_1 & \cdots & a_{1n}+x_n \\ 0 & \vdots & \vdots & \vdots \\ 0 & a_n+x_1 & \cdots & a_{nn}+x_n \end{vmatrix}$$

$$= \begin{vmatrix} 1 & x_1 & \cdots & x_n \\ -1 & a_{11} & \cdots & a_{1n} \\ -1 & \vdots & \vdots & \vdots \\ -1 & a_n & \cdots & a_{nn} \end{vmatrix}$$

$$= det\ A + \sum_{j=1}^{n} x_j(-1)^j\ det\ D_j$$

exercise 5

Definition. The n functions $f_1(x),\ f_2(x),\ \ldots,\ f_n(x)$ are linearly dependent if $\sum_{i=1}^{n} c_i f_i(x) = 0$ for some $c_1,\ c_2,\ \ldots,\ c_n \in \mathbb{R}$ are not all zero.

Definition. If objects in a vector space are functions, we define operations, addition and multiplication by scalars, as follows.

$f(x) = f_1(x) + f_2(x)$, where $f(x)$ is sum of functions $f_1(x)$ and $f_2(x)$.

$f(x) = \lambda g(x)$, where $f(x)$ is multiplication of $g(x)$ by scalar λ.

1. Let $p,\ q,\ r$ and s be polynomials of degree at most 3. Which of the following two conditions is sufficient for the conclusion that the polynomials are linearly dependent (if any)?

   (1) Each of the polynomials has the value 0 at 1.

   (2) Each of the polynomials has the value 1 at 0.

2. Let T be a linear transformation of a vector space $\mathbf{V}$ into itself. Suppose there exists $x \in V$ such that $T^m x = 0,\ T^{m-1}x \neq 0$ for some positive integer m. Show that $x,\ Tx,\ \ldots,\ T^{m-1}x$ are linearly independent.

3. Let $\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n$ be distinct real numbers. Show that the n exponential functions $e^{\alpha_1 t},\ e^{\alpha_2 t},\ \ldots,\ e^{\alpha_n t}$ are linearly independent over the real numbers.

4. Let V be the vector sapce of all continuous real-valued functions defined on the open interval $(-\frac{\pi}{2}, \frac{\pi}{2})$.

   (a) Prove that the set $\{sin\ x, cos\ x, tan\ x, sec\ x\}$ is linearly independent.

   (b) Let $\mathbf{W}$ be the linear space generated by the four trigonometric functions given in (a), and let T be the linear transformation determined on $\mathbf{W}$ into $\mathbf{V}$ by $T(sinx) = sin^2x,\ T(cosx) = cos^2x,\ T(tanx) = tan^2x$ and $T(secx) = sec^2x$. Find a nonzero function in $\mathbf{W}$ that is in the kernel of T.

5. Let $\alpha 1, \alpha 2, \cdots, \alpha n$ be linearly independent vectors in a vector space V. Discuss whether $\alpha 1 + \alpha 2, \alpha 2 + \alpha 3, \cdots, \alpha n - 1 + \alpha n, \alpha n + \alpha 1$ are linearly dependent or linearly independent.

6. Given subspaces H and K of a vector space V, the sum of H and K, writthen as $H + K$, is the set of all vectors in V that can be written as the sum of two vectors, one in H and the other in K; that is, $H + K = w : w = u + v, u \in H, v \in K$.

   (a) Show that $H + K$ is a subspace of V.

   (b) Show that H is a subspace of $H + K$ and K is a subspace of $H + K$.

   (c) Suppose that $u1, \ldots, up$ and $v1, \ldots, vq$ are vectors in V and let $H = Spanu1, \ldots, up$ and $K = Spanv1, \ldots, vq$. Show that $H + K = Spanu1, \ldots, up, v1, \ldots, vq$.

7. Let $\mathbf{M}_{2\times2}$ be the vector space of all 2×2 matrices, and defiine $T : M_{2\times2} \to M_{2\times2}$ by $T(A) = A + A^T$, where $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

   (a) Show that T is a linear transformation.

   (b) Let B be any element of $M_{2\times2}$ such that $B^T = B$. Find an A in $M_{2\times2}$ such that $T(A) = B$.

   (c) Show that the range of T is the set of B in $M_{2\times2}$ with the property $B^T = B$.

   (d) Describe the kernel of T.

思路:

1. 当x=1 时，四个系数和为0，对应行变换加到同一行，产生一个全零行，置底，可得不满秩
2. 把x映射m-1次，得c1为0，以此类推，所有都为0
3. 多项式恒等0
4. （a）代入值，或泰勒展开

   （b）sinx + cosx - （secx-tanx）　　可以变x，也可以变系数（线性变换）（说明，因为他们是linear independent，所以系数非全零a就不是0）
5. 奇偶分析，奇数可知independent，偶数dependent（举反例即可Give a counterexample，法2直接证：a1+an=(a1+a2)-(a2+a3)+(a3+a4)-()+...(an-3+an-2)-(an-2+an-1)+(an-1+an)）
6. 围绕定义入手，subspace满足子集/零向量包含/运算闭合。（第三问，几个向量的张成即他们线性组合出来的空间）
7. （a）定义

   （b）只要找一个A，可以取1/2 B (证明B=T（1/2 B）即可)

   （c）证明range和set of B 等价需要两步:

   第一步证明 到达域（range）中的矩阵（that is to say，像，image）都满足 “自己等于自己的转置”

   第二步证明 满足 “自己等于自己的转置” 的矩阵都在range中（用到2x2的条件求出一个A即可）

$对于任意B = \begin{bmatrix} e & f \\ g & h \end{bmatrix}, 都可以找到原像（preimage）A = \begin{bmatrix} \frac{e}{2} & \frac{g}{2} \\ \frac{g}{2} & \frac{h}{2} \end{bmatrix}$ (这只是A的一个，A的副对角线元素可以调整)

说明 只要满足 “自己等于自己的转置” 的矩阵都在range中（都是image）

由此我们也可以发现，不同的A是有可能映射到同一个B身上

（d）a=d=0, b+c=0 的 A 构成的集合

Group 1 work

12347

Definition. The n functions $f_1(x),\ f_2(x),\ \ldots,\ f_n(x)$ are linearly dependent if $\sum_{i=1}^{n} c_i f_i(x) = 0$ for some $c_1,\ c_2,\ \ldots,\ c_n \in \mathbb{R}$ are not all zero.

Definition. If objects in a vector space are functions, we define operations, addition and multiplication by scalars, as follows.

$f(x) = f_1(x) + f_2(x)$, where $f(x)$ is sum of functions $f_1(x)$ and $f_2(x)$.

$f(x) = \lambda g(x)$, where $f(x)$ is multiplication of $g(x)$ by scalar λ.

1. Let $p,\ q,\ r$ and s be polynomials of degree at most 3. Which of the following two conditions is sufficient for the conclusion that the polynomials are linearly dependent (if any)?

   (1) Each of the polynomials has the value 0 at 1.

   (2) Each of the polynomials has the value 1 at 0.

solution :

$p(x) = a_{11} + a_{12}x + a_{13}x^2 + a_{14}x^3$

$q(x) = a_{21} + a_{22}x + a_{23}x^2 + a_{24}x^3$

$r(x) = a_{31} + a_{32}x + a_{33}x^2 + a_{34}x^3$

$s(x) = a_{41} + a_{42}x + a_{43}x^2 + a_{44}x^3$

Suppose $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$.

$$\begin{aligned} & v_1 p(x) + v_2 q(x) + v_3 r(x) + v_4 s(x) \\ =& v_1(a_{11} + a_{12}x + a_{13}x^2 + a_{14}x^3) + v_2(a_{21} + a_{22}x + a_{23}x^2 + a_{24}x^3) + v_3(a_{31} + a_{32}x + a_{33}x^2 + a_{34}x^3) + v_4(a_{41} + a_{42}x + a_{43}x^2 + a_{44}x^3) \\ =& (v_1a_{11} + v_2a_{21} + v_3a_{31} + v_4a_{41}) + (v_1a_{12} + v_2a_{22} + v_3a_{32} + v_4a_{42})x + (v_1a_{13} + v_2a_{23} + v_3a_{33} + v_4a_{43})x^2 + (v_1a_{14} + v_2a_{24} + v_3a_{34} + v_4 \end{aligned}$$

the polynomials are linearly dependent

$\Leftrightarrow v_1 p(x) + v_2 q(x) + v_3 r(x) + v_4 s(x) = 0$ for some $v_1,\ v_2,\ v_3,\ v_4$ are not all zero.

$\Leftrightarrow A^T\mathbf{v} = \mathbf{0}$ has non-trivial solution, where $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix}$.

(1) Each of the polynomials has the value 0 at 1.

$\Rightarrow A\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \mathbf{0}$

$\Rightarrow A\mathbf{x} = \mathbf{0}$ has non-trivial solution.

$\Rightarrow A$ is not invertible.

$\Rightarrow A^T$ is not invertible.

$\Rightarrow A^T\mathbf{v} = \mathbf{0}$ has non-trivial solution.

$\Rightarrow$ The polynomials are linearly dependent.

(2) Each of the polynomials has the value 1 at 0.

Give a counterexample:

$p(x) = 1$

$q(x) = 1 + x$

$r(x) = 1 + x + x^2$

$s(x) = 1 + x + x^2 + x^3$

$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$

$det A = 1 \neq 0$

$\Rightarrow A$ is invertible.

$\Rightarrow A^T$ is invertible.

$\Rightarrow A^T \mathbf{v} = \mathbf{0}$ has only trivial solution.

$\Rightarrow$ The polynomials are **not** linearly dependent.

2. Let T be a linear transformation of a vector space $\mathbf{V}$ into itself. Suppose there exists $x \in V$ such that $T^m x = 0,\ T^{m-1} x \neq 0$ for some positive integer m. Show that $x,\ Tx,\ \ldots,\ T^{m-1}x$ are linearly independent.

$\textbf{solution}$:

Suppose $c_1 x + c_2 Tx + \cdots + c_m T^m x = 0$, then

$T^{m-1}(c_1 x + c_2 Tx + \cdots + c_m T^m x) = T^{m-1} 0 = 0$

$T^{m-1} x \neq 0,\ c_1 T^{m-1} x = 0 \Rightarrow c_1 = 0$. Then

$c_2 Tx + \cdots + c_m T^m x = 0$

$T^{m-2}(c_2 Tx + \cdots + c_m T^m x) = T^{m-2} 0 = 0$

$T^{m-1} x \neq 0,\ c_2 T^{m-1} x = 0 \Rightarrow c_2 = 0$.

And so on, we can find that $c_1 = c_2 = \cdots = c_m = 0 \Rightarrow x,\ Tx,\ \ldots,\ T^{m-1}x$ are linearly independent.

3. Let $\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n$ be distinct real numbers. Show that the n exponential functions $e^{\alpha_1 t},\ e^{\alpha_2 t},\ \ldots,\ e^{\alpha_n t}$ are linearly independent over the real numbers.

$\textbf{solution}$:

Suppose $\alpha_1 > \alpha_2 > \cdots > \alpha_n$.

Let $c_1 e^{\alpha_1 t} + c_2 e^{\alpha_2 t} + \cdots + c_n e^{\alpha_n t} = 0$, then

$c_1 + c_2 e^{-(\alpha_1 - \alpha_2)t} + \cdots + c_n e^{-(\alpha_1 - \alpha_n)t} = 0$

$t \to \infty$, $c_1 = 0$. Then,

$c_2 e^{\alpha_2 t} + \cdots + c_n e^{\alpha_n t} = 0$

$c_2 + c_3 e^{-(\alpha_2 - \alpha_3)t} + \cdots + c_n e^{-(\alpha_2 - \alpha_n)t} = 0$

$t \to \infty$, $c_2 = 0$. And so on, we can find $c_1 = c_2 = \cdots = c_n = 0 \Rightarrow e^{\alpha_1 t},\ e^{\alpha_2 t},\ \ldots,\ e^{\alpha_n t}$ are linearly independent.

4. Let V be the vector sapce of all continuous real-valued functions defined on the open interval $(-\frac{\pi}{2}, \frac{\pi}{2})$.

(a) Prove that the set $\{sin\ x, cos\ x, tan\ x, sec\ x\}$ is linearly independent.

(b) Let $\mathbf{W}$ be the linear space generated by the four trigonometric functions given in (a), and let T be the linear transformation determined on $\mathbf{W}$ into $\mathbf{V}$ by $T(sinx) = sin^2x,\ T(cosx) = cos^2x,\ T(tanx) = tan^2x$ and $T(secx) = sec^2x$. Find a nonzero function in $\mathbf{W}$ that is in the kernel of T.

solution :

(a) Let $c_1sin\ x + c_2cos\ x + c_3tan\ x + c_4sec\ x = 0$, x takes the values $\frac{\pi}{6}, -\frac{\pi}{6}, \frac{\pi}{4}, -\frac{\pi}{4}$ respectively.

$$\begin{cases} \frac{1}{2}c_1 + \frac{\sqrt{3}}{2}c_2 + \frac{\sqrt{3}}{3}c_3 + \frac{2}{\sqrt{3}}c_4 = 0 \\ -\frac{1}{2}c_1 + \frac{\sqrt{3}}{2}c_2 - \frac{\sqrt{3}}{3}c_3 + \frac{2}{\sqrt{3}}c_4 = 0 \\ \frac{\sqrt{2}}{2}c_1 + \frac{\sqrt{2}}{2}c_2 + c_3 + \sqrt{2}\,c_4 = 0 \\ -\frac{\sqrt{2}}{2}c_1 + \frac{\sqrt{2}}{2}c_2 - c_3 + \sqrt{2}\,c_4 = 0 \end{cases} \Rightarrow \begin{cases} c_1 = 0 \\ c_2 = 0 \\ c_3 = 0 \\ c_4 = 0 \end{cases} \Rightarrow \text{the set } \{sin\ x, cos\ x, tan\ x, sec\ x\} \text{ is linearly independent.}$$

(b) Because$(sin^2x + cos^2x) - (sec^2x - tan^2x) = 0$

Hence we can find a function $(sin\ x + cos\ x + tan\ x - sec\ x)$ in $\mathbf{W}$ that is in the kernel of T.

7. Let $\mathbf{M}_{2\times2}$ be the vector space of all 2×2 matrices, and defiine $T : M_{2\times2} \to M_{2\times2}$ by $T(A) = A + A^T$, where $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

(a) Show that T is a linear transformation.

(b) Let B be any element of $M_{2\times2}$ such that $B^T = B$. Find an A in $M_{2\times2}$ such that $T(A) = B$.

(c) Show that the range of T is the set of B in $M_{2\times2}$ with the property $B^T = B$.

(d) Describe the kernel of T.

solution :

(a)

$\forall A, B \in M_{2\times2}$,

$T(A+B) = (A+B) + (A+B)^T = (A + A^T) + (B + B^T) = T(A) + T(B)$

$\forall A \in M_{2\times2},\ c \in \mathbb{R}$,

$T(cA) = cA + (cA)^T = c(A + A^T) = cT(A)$

Hence T is a linear transformation.

(b)

$T(\frac{B}{2}) = \frac{B}{2} + (\frac{B}{2})^T = B$,

Then we find an $A = \frac{B}{2}$ in $M_{2\times2}$ such that $T(A) = B$.

(c)

$\forall A \in M_{2\times2},\ (T(A))^T = (A + A^T)^T = A + A^T = T(A)$.

the range of $T \subseteq$ the set of B in $M_{2\times2}$ with the property $B^T = B$

Suppose $B = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$,

$\forall B$ in $M_{2\times 2}$ with the property $B^T = B,\ f = g \Rightarrow B = \begin{bmatrix} e & f \\ f & h \end{bmatrix}$.

We can always find an $A = \begin{bmatrix} \frac{e}{2} & \frac{f}{2} \\ \frac{f}{2} & \frac{h}{2} \end{bmatrix}$ such that $T(A) = B$.

the range of $T \supseteq$ the set of B in $M_{2\times 2}$ with the property $B^T = B$

Hence the range of T is the set of B in $M_{2\times 2}$ with the property $B^T = B$.

(d) The kernel of $T = \{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2} : a = d = 0,\ b + c = 0\}$.

exercise 6

1. Assume the matrix A is $m \times n$ and the matrix B is $n \times p$. Prove the following useful properties of rank.

   (a). $\operatorname{rank} AB \leq \operatorname{rank} A$.

   (b). If $AB = 0$, then $\operatorname{rank} A + \operatorname{rank} B \leq n$.

2. Let A be an $m \times s$ matrix. Assume the equation $Ax = \beta$ has solution for all non-zero vector $\beta \in \mathbb{R}^m$. Prove $\operatorname{rank} A = m$.

   solution:
   $\forall\ \beta \in \mathbb{R}^m\ \beta \in Col(A)$ 零矢量单独算（显然0在Rm里 trivial）

3. Assume the matrix A is $s \times n$, matrix B is $n \times m$, and $\operatorname{rank} AB = \operatorname{rank} B$.

   (a). Prove that $\operatorname{Nul} AB = \operatorname{Nul} B$.

   (b). For all $m \times l$ matrix C, do we always have $\operatorname{Nul} ABC = \operatorname{Nul} BC$? Moreover, do we always have $\operatorname{rank} ABC = \operatorname{rank} BC$?

4. Suppose V is an n-dimensional vector space. Let $W \subset V$ be a subspace of dimension $r < n$. Show that

$$W = \cap\{U \mid U \text{ is an } (n-1)- \text{ dimensional subspace of } V \text{ and } W \subset U\}.$$

(Hint: Denote Y as the intersection given above. You may try to prove $Y \subset W$ by contradiction.)

5. Let $A = \begin{pmatrix} 1 & -1 & 0 & -1 & -2 \\ -1 & 2 & 1 & 3 & 6 \\ 0 & 1 & 1 & 2 & 4 \\ 0 & -1 & -1 & 1 & 2 \end{pmatrix}$. Denote $\mathbb{R}^{5\times 2}$ as the vector space of all 5×2 matrix with elements in $\mathbb{R}$. Define W as

$$W := \left\{B \in \mathbb{R}^{5\times 2}; AB = 0\right\}.$$

(a). Prove $\dim \ker A = 2$ and find a basis of it.

(b). Prove W is a subspace of $\mathbb{R}^{5\times 2}$ and calculate $\dim W$.

(Hint: You can represent W by the result in (a).)

6. If $\mathbf{V}$ is an n-dimensional vector space over $\mathbb{R}$, and $f, g : \mathbf{V} \longrightarrow \mathbb{R}$ are two linear functions on $\mathbf{V}$. Assume ker $f = \ker g$, prove there exists some non-zero constant c such that $g = cf$.

(Hint: The kernel for a function $f : \mathbf{V} \longrightarrow \mathbb{R}$ is the set $\{x \in \mathbf{V} \mid f(x) = 0\}$. And you may use the following form of the Rank Theorem.)

Theorem. If $T : V \to W$ is a linear map, then $\dim(\operatorname{Im}(T)) + \dim(\ker(T)) = \dim(V)$, where $\operatorname{Im}(T)$ is the range of T.

7. Let $A = (a_{ij})_{i,j=1}^n$ be a real $n \times n$ matrix with nonnegative entries such that

$$\sum_{j=1}^n a_{ij} = 1, \quad 1 \leq i \leq n$$

Prove that no eigenvalue of A has absolute value greater than 1 .

Group 1 work

1,3,4,5,6

1. Assume the matrix A is $m \times n$ and the matrix B is $n \times p$. Prove the following useful properties of rank.

(a). $\operatorname{rank} AB \leq \operatorname{rank} A$.

(b). If $AB = 0$, then $\operatorname{rank} A + \operatorname{rank} B \leq n$.

Solution:

(a) To prove $\operatorname{rank} AB \leq \operatorname{rank} A$, we just need to prove $col(AB) \subseteq col(A)$.

$$AB = [\, A\mathbf{b_1} \quad A\mathbf{b_2} \quad \cdots \quad A\mathbf{b_p} \,]$$
$$= \left[[\, \mathbf{a_1} \quad \mathbf{a_2} \quad \cdots \quad \mathbf{a_n} \,] \begin{bmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{n1} \end{bmatrix} \quad [\, \mathbf{a_1} \quad \mathbf{a_2} \quad \cdots \quad \mathbf{a_n} \,] \begin{bmatrix} b_{12} \\ b_{22} \\ \vdots \\ b_{n2} \end{bmatrix} \quad \cdots \quad [\, \mathbf{a_1} \quad \mathbf{a_2} \quad \cdots \quad \mathbf{a_n} \,] \begin{bmatrix} b_{1p} \\ b_{2p} \\ \vdots \\ b_{np} \end{bmatrix} \right]$$

$$\forall\, \mathbf{v} \in col(AB),\ \mathbf{v} = c_1 [\, \mathbf{a_1} \quad \mathbf{a_2} \quad \cdots \quad \mathbf{a_n} \,] \begin{bmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{n1} \end{bmatrix} + c_2 [\, \mathbf{a_1} \quad \mathbf{a_2} \quad \cdots \quad \mathbf{a_n} \,] \begin{bmatrix} b_{12} \\ b_{22} \\ \vdots \\ b_{n2} \end{bmatrix} + \cdots + c_p [\, \mathbf{a_1} \quad \mathbf{a_2} \quad \cdots \quad \mathbf{a_n} \,] \begin{bmatrix} b_{1p} \\ b_{2p} \\ \vdots \\ b_{np} \end{bmatrix}$$
$$= c_1'\mathbf{a_1} + c_2'\mathbf{a_2} + \cdots + c_n'\mathbf{a_n}$$

$$\Rightarrow \mathbf{v} \in col(A) \Rightarrow col(AB) \subseteq col(A)$$

(b)

$AB = 0 \Rightarrow A [\, \mathbf{b_1} \quad \mathbf{b_2} \quad \cdots \quad \mathbf{b_p} \,] = 0.$

$\forall\, \mathbf{b} \in Col\ B,\ \mathbf{b} = c_1\mathbf{b_1} + c_2\mathbf{b_2} + \cdots + c_p\mathbf{b_p}.$

$A\mathbf{b} = A(c_1\mathbf{b_1} + c_2\mathbf{b_2} + \cdots + c_p\mathbf{b_p}) = \mathbf{0}$

$\Rightarrow \mathbf{b} \in Nul\ A$

$\Rightarrow Col\ B \subseteq Nul\ A.$

$$\begin{cases} rank\ A + dim\ Nul\ A = n \\ rank\ B = dim\ Col\ B \leq dim\ Nul\ A \end{cases} \Rightarrow \ \text{rank}\, A + \text{rank}\, B \leq n.$$

3. Assume the matrix A is $s \times n$, matrix B is $n \times m$, and $\text{rank}\, AB = \text{rank}\, B$.

(a). Prove that $\text{Nul}\, AB = \text{Nul}\, B$.

(b). For all $m \times l$ matrix C, do we always have $\text{Nul}\, ABC = \text{Nul}\, BC$? Moreover, do we always have $\text{rank}\, ABC = \text{rank}\, BC$?

solution:

(a)

$\forall\, \mathbf{x} \in \mathbb{R}^m;\ B\mathbf{x} = \mathbf{0},$

$(AB)\mathbf{x} = A(B\mathbf{x}) = \mathbf{0}.$

$\Rightarrow Nul\ B \subseteq Nul\ AB.$

By rank theorem $\Rightarrow \begin{cases} rank\ B + dim\ Nul\ B = n \\ rank\ AB + dim\ Nul\ AB = n \end{cases}$,

$\text{rank}\, AB = \text{rank}\, B \Rightarrow dim\ Nul\ AB = dim\ Nul\ B \Rightarrow \text{Nul}\, AB = \text{Nul}\, B.$

(b) Yes.

$\forall\, \mathbf{x} \in \mathbb{R}^m;\ (BC)\mathbf{x} = \mathbf{0},$

$(ABC)\mathbf{x} = A((BC)\mathbf{x}) = \mathbf{0}.$

$\Rightarrow Nul\ BC \subseteq Nul\ ABC.$

$\forall\, \mathbf{x} \in \mathbb{R}^l;\ \mathbf{x} \in Nul\ ABC,$

$AB(C\mathbf{x}) = (ABC)\mathbf{x} = \mathbf{0}.$

By part (a),

$C\mathbf{x} \in Nul\ AB \Leftrightarrow C\mathbf{x} \in Nul\ B.$

$(BC)\mathbf{x} = B(C\mathbf{x}) = \mathbf{0} \Rightarrow \mathbf{x} \in Nul\ BC$

$\Rightarrow Nul\ ABC \subseteq Nul\ BC.$

Hence $\text{Nul}\, ABC = \text{Nul}\, BC \Rightarrow dim\ Nul\ ABC = dim\ Nul\ BC.$

By rank theorem $\Rightarrow \begin{cases} rank\ BC + dim\ Nul\ BC = l \\ rank\ ABC + dim\ Nul\ ABC = l' \end{cases}$

$dim\ Nul\ ABC = dim\ Nul\ BC \Rightarrow \text{rank}\, ABC = \text{rank}\, BC.$

4. Suppose V is an n-dimensional vector space. Let $W \subset V$ be a subspace of dimension $r < n$. Show that

$$W = \cap\{U \mid U \text{ is an } (n-1)- \text{ dimensional subspace of } V \text{ and } W \subset U\}.$$

(Hint: Denote Y as the intersection given above. You may try to prove $Y \subset W$ by contradiction.)

solution:

Let $Y = \cap\{U \mid U \text{ is an } (n-1)- \text{ dimensional subspace of } V \text{ and } W \subset U\}.$

$\forall\, w \in W,\ w \in Y \Rightarrow W \subseteq Y.$

To prove $Y \subseteq W$, assume $\exists\, \beta \in Y, \beta \notin W.$

Because $dim\ W = r$, we can find a basis for W $\{\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_r\}$.

$\beta \notin W \Rightarrow \{\beta,\ \alpha_1,\ \alpha_2,\ \cdots,\ \alpha_r\}$ is a linearly independent set, with $Span\{\beta,\ \alpha_1,\ \alpha_2,\ \cdots,\ \alpha_r\} \subseteq V$.

Then we can find another $n-(r+1)$ vector $\alpha_{r+1},\ \alpha_{r+2},\ \cdots,\ \alpha_{n-1}$, such that $\{\beta,\ \alpha_1,\ \alpha_2,\ \cdots,\ \alpha_r,\ \alpha_{r+1},\ \alpha_{r+2},\ \cdots,\ \alpha_{n-1}\}$ is a linearly independent set and $Span\{\beta,\ \alpha_1,\ \alpha_2,\ \cdots,\ \alpha_r,\ \alpha_{r+1},\ \alpha_{r+2},\ \cdots,\ \alpha_{n-1}\} = V$.

Let $Z = Span\{\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_{n-1}\}$, then

$$\begin{cases} dim\ Z = n-1 \\ W \subseteq Z \end{cases} \Rightarrow Y \subseteq Z \Rightarrow \beta \subseteq Z.$$

But $\{\beta,\ \alpha_1,\ \alpha_2,\ \cdots,\ \alpha_r,\ \alpha_{r+1},\ \alpha_{r+2},\ \cdots,\ \alpha_{n-1}\}$ is a linearly independent set $\Rightarrow \beta \nsubseteq Z$, which lead to contradiction.

Hence, $\forall\, \beta \in Y, \beta \in W \Rightarrow Y \subseteq W.$

Then $W = Y$.

5. Let $A = \begin{pmatrix} 1 & -1 & 0 & -1 & -2 \\ -1 & 2 & 1 & 3 & 6 \\ 0 & 1 & 1 & 2 & 4 \\ 0 & -1 & -1 & 1 & 2 \end{pmatrix}$. Denote $\mathbb{R}^{5\times 2}$ as the vector space of all 5×2 matrix with elements in $\mathbb{R}$. Define W as

$$W := \left\{B \in \mathbb{R}^{5\times 2}; AB = 0\right\}.$$

(a). Prove $\dim \ker A = 2$ and find a basis of it.

(b). Prove W is a subspace of $\mathbb{R}^{5\times 2}$ and calculate $\dim W$.

(Hint: You can represent W by the result in (a).)

solution:

(a) Row reduce the augmented matrix to reduced echelon form.

$$A = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

There are 3 pivot columns and 2 free variables.

$\Rightarrow$ $dim\ ker\ A = 2.$

$$\begin{aligned} x_4 &= -2x_5 \\ x_2 &= -x_3 \\ x_1 = x_2 &= -x_3 \end{aligned}$$

$$ker A = \begin{bmatrix} -x_3 \\ -x_3 \\ x_3 \\ -2x_5 \\ x_5 \end{bmatrix} = x_3 \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 0 \\ 0 \\ 0 \\ -2 \\ 1 \end{bmatrix}$$

$\{ \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ -2 \\ 1 \end{bmatrix} \}$ is one of the basis.

(b) $\forall\, U, V \in W \Rightarrow\ U + V \in W$

$\forall\, U \in W, r \in \mathbb{R} \Rightarrow\ rU \in W$

$\Rightarrow$ W is a subspace of $\mathbb{R}^{5\times 2}$.

$\forall\, B \in \mathbb{R}^{5\times 2};\, AB = 0 \Rightarrow \mathbf{b_1} \in ker\ A;\ \mathbf{b_2} \in ker\ A.$

$$\mathbf{b_1} = \lambda_1 \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 \\ 0 \\ 0 \\ -2 \\ 1 \end{bmatrix},\ \mathbf{b_2} = \lambda_3 \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \lambda_4 \begin{bmatrix} 0 \\ 0 \\ 0 \\ -2 \\ 1 \end{bmatrix}.$$

$$B = \lambda_1 \begin{bmatrix} -1 & 0 \\ -1 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ -2 & 0 \\ 1 & 0 \end{bmatrix} + \lambda_3 \begin{bmatrix} 0 & -1 \\ 0 & -1 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} + \lambda_4 \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & -2 \\ 0 & 1 \end{bmatrix}.$$

Let $\lambda_1 \begin{bmatrix} -1 & 0 \\ -1 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ -2 & 0 \\ 1 & 0 \end{bmatrix} + \lambda_3 \begin{bmatrix} 0 & -1 \\ 0 & -1 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} + \lambda_4 \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & -2 \\ 0 & 1 \end{bmatrix} = 0,$

$$\begin{cases} \lambda_1 \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 \\ 0 \\ 0 \\ -2 \\ 1 \end{bmatrix} = 0 \\ \lambda_3 \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \lambda_4 \begin{bmatrix} 0 \\ 0 \\ 0 \\ -2 \\ 1 \end{bmatrix} = 0 \end{cases} \Rightarrow \lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0.$$

$$\Rightarrow \left\{ \begin{bmatrix} -1 & 0 \\ -1 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ -2 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 0 & -1 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & -2 \\ 0 & 1 \end{bmatrix} \right\}$$ is a linearly independent set as well as a basis of W.

$\Rightarrow\ dim\ W = 4.$

6. If $\mathbf{V}$ is an n-dimensional vector space over $\mathbb{R}$, and $f, g : \mathbf{V} \longrightarrow \mathbb{R}$ are two linear functions on $\mathbf{V}$. Assume $\ker f = \ker g$, prove there exists some non-zero constant c such that $g = cf$.

(Hint: The kernel for a function $f : \mathbf{V} \longrightarrow \mathbb{R}$ is the set $\{x \in \mathbf{V} \mid f(x) = 0\}$. And you may use the following form of the Rank Theorem.)

Theorem. If $T : V \to W$ is a linear map, then $\dim(\operatorname{Im}(T)) + \dim(\ker(T)) = \dim(V)$, where $\operatorname{Im}(T)$ is the range of T.

solution:

$$\begin{cases} \dim(\operatorname{Im}(f)) + \dim(\ker(f)) = \dim(V) \\ \dim(\operatorname{Im}(g)) + \dim(\ker(g)) = \dim(V) \Rightarrow \dim(\operatorname{Im}(f)) = \dim(\operatorname{Im}(g)) = 0\ or\ 1. \\ \ker f = \ker g \end{cases}$$

(1) $\dim(\operatorname{Im}(f)) = \dim(\operatorname{Im}(g)) = 0.$

$\dim(\operatorname{Im}(f)) = 0 \Rightarrow dim(ker\ f) = dim\ V.$

$ker\ f \subseteq V \Rightarrow ker\ f = V.$

Similarly, $ker\ g = V.$

$\forall\ \mathbf{w} \in V,\ f(\mathbf{w}) = g(\mathbf{w}) = \mathbf{0}.$

$\Rightarrow$ There exists some non-zero constant $c = 1,\ 2, \cdots$ such that $g = cf$.

(2) $\dim(\operatorname{Im}(f)) = \dim(\operatorname{Im}(g)) = 1.$

$\dim(\operatorname{Im}(f)) = 1 \Rightarrow dim(ker\ f) = n - 1.$

Construct a basis for $ker\ f$ and $ker\ g$: $\{\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_{n-1}\}$.

Because $ker\ f \subset V$, we can always find an vector α_n such that $\{\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_{n-1},\ \alpha_n\}$ is a basis of V.

$\forall\ \mathbf{w} \in V,\ \mathbf{w} = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n.$

$$\begin{cases} f(\mathbf{w}) = k_n f(\alpha_n) \\ g(\mathbf{w}) = k_n g(\alpha_n) \end{cases}$$

$\Rightarrow$ There exists non-zero constant $c = \frac{g(\alpha_n)}{f(\alpha_n)}$ such that $g = cf$.

exercise 7

1. Let A_n be the $n \times n$ matrix whose entries a_{jk} are given by

$$a_{jk} = \begin{cases} 1 & \text{if } |j-k| = 1 \\ 0 & \text{otherwise} \end{cases}$$

Prove that the eigenvalues of A are symmetric with respect to the origin, that is, if λ is an eigenvalue of A, then $-\lambda$ is also an eigenvalue of A.

(Hint: Find the recurrence relation of the characteristic polynomial of A_n.)

2. If $p(t) = c_0 + c_1 t + c_2 t^2 + \cdots + c_n t^n$, define $p(A)$ as the matrix formed by replacing each power of t in $p(t)$ by the corresponding power of A (with $A^0 = I$). That is,

$$p(A) = c_0 I + c_1 A + c_2 A^2 + \cdots c_n A^n$$

(a) Show that if λ is an eigenvalue of A, then one eigenvalue of $p(A)$ is $p(\lambda)$.

(b) Suppose A is diagonalizable. Show that $p(A)$ is diagonalizable.

(c) Suppose A is diagonalizable and $p(t)$ is the characteristic polynomial of A. Show that $p(A)$ is the zero matrix. This fact is called the Cayley-Hamilton theorem.

(Hint for (c) : Use the result in (b) to show that the diagonal entries of $p(D)$ are zero.)

3. Suppose A is an $n \times n$ matrix, λ_1 and λ_2 are two distinct eigenvalues of A corresponding to eigenvectors X_1 and X_2. Prove $X_1 + X_2$ is not the eigenvector of A.

4. For every n, if the eigenvalues of a symmetric $n \times n$ matrix with entries 0 and 1 are all positive, prove that such matrix must be the identity matrix.

Fact. The following two are equivalent.

(i) All eigenvalues of matrix A are positive.

(ii) $\forall \mathbf{v} \in \mathbb{R}^n$ and $\mathbf{v} \neq 0, \mathbf{v}^T A \mathbf{v} > 0.$

Group 1 work

1,2,3

1. Let A_n be the $n \times n$ matrix whose entries a_{jk} are given by

$$a_{jk} = \begin{cases} 1 & \text{if } |j-k| = 1 \\ 0 & \text{otherwise} \end{cases}$$

Prove that the eigenvalues of A are symmetric with respect to the origin, that is, if λ is an eigenvalue of A, then $-\lambda$ is also an eigenvalue of A.

(Hint: Find the recurrence relation of the characteristic polynomial of A_n.)

Solution:

$$det(A_n - \lambda I) = \begin{vmatrix} -\lambda & 1 & 0 & \cdots & 0 & 0 \\ 1 & -\lambda & 1 & \cdots & 0 & 0 \\ 0 & 1 & -\lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\lambda & 1 \\ 0 & 0 & 0 & \cdots & 1 & -\lambda \end{vmatrix}$$

$$= -\lambda det(A_{n-1} - \lambda I) - \begin{vmatrix} 1 & 0 & \cdots & 0 & 0 \\ 1 & -\lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -\lambda & 1 \\ 0 & 0 & \cdots & 1 & -\lambda \end{vmatrix}$$

$$= -\lambda det(A_{n-1} - \lambda I) - det(A_{n-2} - \lambda I)$$

If $det(A_{n-2} - \lambda I)$ is an odd function of λ and $det(A_{n-1} - \lambda I)$ is an even function of λ, then

$$\begin{aligned} det(A_n - \lambda I) &= -\lambda \cdot (even\ function) - (odd\ function) \\ &= (odd\ function) - (odd\ function) \\ &= odd\ function \end{aligned}$$

If $det(A_{n-2} - \lambda I)$ is an even function of λ and $det(A_{n-1} - \lambda I)$ is an odd function of λ, then

$$\begin{aligned} det(A_n - \lambda I) &= -\lambda \cdot (odd\ function) - (even\ function) \\ &= (even\ function) - (even\ function) \\ &= even\ function \end{aligned}$$

$det(A_1 - \lambda I) = -\lambda$, which is odd, and

$det(A_2 - \lambda I) = \lambda^2 - 1$, which is even. Then the odd and even functions appear alternately.

$f(\lambda) = det(A_n - \lambda I)$ is either odd or even.

$f(\lambda) = 0 \Rightarrow f(-\lambda) = 0$.

2. If $p(t) = c_0 + c_1 t + c_2 t^2 + \cdots + c_n t^n$, define $p(A)$ as the matrix formed by replacing each power of t in $p(t)$ by the corresponding power of A (with $A^0 = I$). That is,

$$p(A) = c_0 I + c_1 A + c_2 A^2 + \cdots c_n A^n$$

(a) Show that if λ is an eigenvalue of A, then one eigenvalue of $p(A)$ is $p(\lambda)$.

(b) Suppose A is diagonalizable. Show that $p(A)$ is diagonalizable.

(c) Suppose A is diagonalizable and $p(t)$ is the characteristic polynomial of A. Show that $p(A)$ is the zero matrix. This fact is called the Cayley-Hamilton theorem.

(Hint for (c) : Use the result in (b) to show that the diagonal entries of $p(D)$ are zero.)

solution:

(a) If λ is an eigenvalue of A, then

$A\mathbf{x} = \lambda\mathbf{x}$, where λ is the eigenvalue and the nonzero vector $\mathbf{x}$ is the eigenvector.

$A^n\mathbf{x} = \lambda A^{n-1}\mathbf{x} = \cdots = \lambda^n\mathbf{x}$, then

$$\begin{aligned} p(A)\mathbf{x} &= (c_0 I + c_1 A + c_2 A^2 + \cdots c_n A^n)\mathbf{x} \\ &= c_0 I\mathbf{x} + c_1 A\mathbf{x} + c_2 A^2\mathbf{x} + \cdots c_n A^n\mathbf{x} \\ &= c_0\mathbf{x} + c_1\lambda\mathbf{x} + c_2\lambda^2\mathbf{x} + \cdots c_n\lambda^n\mathbf{x} \\ &= p(\lambda)\mathbf{x} \end{aligned}$$

$\Rightarrow$ one eigenvalue of $p(A)$ is $p(\lambda)$.

(b) A is diagonalizable $\Rightarrow \exists \begin{cases} invertible\ matrix\ Q \\ diagonal\ matrix\ D \end{cases}$ such that

$A = QDQ^{-1}$. Then, for any $n \in \mathbb{N}^*$,

$A^n = (QDQ^{-1})^n = QD^nQ^{-1}$ is also diagonalizable.

$$\begin{aligned} p(A) &= c_0 I + c_1 A + c_2 A^2 + \cdots c_n A^n \\ &= c_0 I + c_1(QDQ^{-1}) + c_2(QDQ^{-1})^2 + \cdots c_n(QDQ^{-1})^n \\ &= Qc_0IQ^{-1} + Qc_1DQ^{-1} + Qc_2D^2Q^{-1} + \cdots Qc_nD^nQ^{-1} \\ &= Qp(D)Q^{-1} \end{aligned}$$

where $p(D) = c_0 I + c_1 D + c_2 D^2 + \cdots c_n D^n$ is also a diagonal matrix.

$\Rightarrow p(A)$ is diagonalizable.

(c) By (b),

$$p(D) = \begin{bmatrix} p(\lambda_1) & 0 & \cdots & 0 \\ 0 & p(\lambda_2) & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & p(\lambda_n) \end{bmatrix}$$

$p(t)$ is the characteristic polynomial of $A \Rightarrow p(\lambda_1) = p(\lambda_2) = \cdots = p(\lambda_n) = 0$.

$\Rightarrow p(A) = Qp(D)Q^{-1}$ is the zero matrix.

3. Suppose A is an $n \times n$ matrix, λ_1 and λ_2 are two distinct eigenvalues of A corresponding to eigenvectors X_1 and X_2. Prove $X_1 + X_2$ is not the eigenvector of A.

solution:

Assume $X_1 + X_2$ is the eigenvector of A, and

$A(X_1 + X_2) = \lambda_0(X_1 + X_2)$. Then,

$A(X_1 + X_2) = AX_1 + AX_2 = \lambda_1 X_1 + \lambda_2 X_2$.

$\Rightarrow (\lambda_1 - \lambda_0)X_1 + (\lambda_2 - \lambda_0)X_2 = \mathbf{0}$.

λ_1 and λ_2 are distinct $\Rightarrow$ X_1 and X_2 are linearly independent.

$\Rightarrow \lambda_1 = \lambda_2 = \lambda_0$, which is impossible.

Hence $X_1 + X_2$ is not the eigenvector of A.

exercise 8

Definition. Consider a linear mapping $T : \mathbb{R}^n \to \mathbb{R}^n$. A T-invariant subspace W has the property that any vector $v \in W$ is transformed by T into a vector also contained in W, i.e., $v \in W \Longrightarrow T(v) \in W$.

Definition. Consider any linear mapping $T : V \to M$ for two vector spaces V and M. If $T(w) = \lambda w$ for some $\lambda \in \mathfrak{R}$ and a nonzero element $w \in V$, we say λ is an eigenvalue and w is the associated eigenvector for T. Denote S_λ the corresponding eigenspace, which consists of all eigenvectors associated with λ.

Definition. An invertible matrix A is orthogonal if $A^{-1} = A^\top$.

1. Let λ_1, λ_2 be two different eigenvalues of the symmetric matrix A. And v_1, v_2 are the associated eigenvectors with λ_1, λ_2 respectively. Show that $v_1^T v_2 = 0$.

2. Suppose A and M are $n \times n$ real matrices with nonzero eigenvalues. A is invertible and $AMA^{-1} = M^2$. Prove that for any nonzero eigenvalues of M, there exists some $m \in \mathbb{Z}_{++}$such that the eigenvalue is a root of $\alpha^m = 1$. Moreover, there exists some $n \in \mathbb{Z}_{++}$such that any nonzero eigenvalue of M is the root of $\alpha^n = 1$.

3. You can use the following observation without proof: For any linear transformation T, if there exists a T-invariant subspace V, then T must have an eigenvector in V.

Let A and B are real $n \times n$ matrices such that both are with real eigenvalues; moreover, $AB = BA$. Prove that A and B have a common eigenvector.

4. Let V be the vector space of sequences (a_n) of real numbers. The shift operator $S : V \to V$ is defined by

$$S((a_1, a_2, a_3, \ldots)) = (a_2, a_3, a_4, \ldots).$$

(a) Given an eigenvalue of S being λ, find the corresponding eigenvectors.

(b) Show that the subspace W consisting of the sequences (x_n) with $x_{n+2} = x_{n+1} + x_n$ is a twodimensional, S-invariant subspace of V and exhibit an explicit basis for W.

(c) Find an explicit formula for the Fibonacci number f_n, where $f_2 = f_1 = 1, f_{n+2} = f_{n+1} + f_n$ for $n \geq 1$.

5. Find a 2×2 orthogonal matrix with no real eigenvalues. And show that, for any positive integer n that is divisible by 2 , there exists an $n \times n$ orthogonal matrix that has no real eigenvalues.

solution:

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

$A^{-1} = A^T$

$det(A^{-1} - \lambda I) = det(A^T - \lambda I) = det((A - \lambda I)^T) = det(A - \lambda I)$

$A\mathbf{x} = \lambda \mathbf{x} \Rightarrow A^{-1}\mathbf{x} = \lambda \mathbf{x}$

$AA^{-1}\mathbf{x} = \mathbf{x} = A(\lambda \mathbf{x}) = \lambda^2 \mathbf{x}$

$\Rightarrow \lambda = 1$ or $\lambda = -1$

$det(A^T) = det(A^{-1}) = det(A)$

$det(A)det(A^{-1}) = 1$

$\Rightarrow det(A) = 1$ or -1

$$\begin{bmatrix} 0 & 0 & \cdots & 0 & -1 \\ 0 & \vdots & \cdots & 1 & 0 \\ \vdots & 0 & \ddots & 0 & \vdots \\ 0 & -1 & \cdots & \vdots & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

6. Let A be a $n \times n$ matrix such that: 1) all eigenvalues are positive; 2) A is diagonalizable with orthogonal matrix. Let $y \in \mathbb{R}^n, y \neq 0$. Prove that the limit

$$\lim_{m \to \infty} \frac{y^\top A^{m+1} y}{y^\top A^m y}$$

exists and is an eigenvalue of A.

7. Let k be real, n be an integer ≥ 2, and let $A = (a_{ij})$ be the $n \times n$ matrix such that all diagonal entries $a_{ii} = k$, all entries $a_{i,i\pm1}$ immediately above or below the diagonal equal 1 , and all other entries equal 0. For example, if $n = 5$,

$$A = \begin{pmatrix} k & 1 & 0 & 0 & 0 \\ 1 & k & 1 & 0 & 0 \\ 0 & 1 & k & 1 & 0 \\ 0 & 0 & 1 & k & 1 \\ 0 & 0 & 0 & 1 & k \end{pmatrix}$$

Moreover, we choose k such that all eigenvalues for A are positive. Let $\lambda_{\min}$ and $\lambda_{\max}$ denote the smallest and largest eigenvalues of A, respectively. Use the following fact and show that $\lambda_{\min} \leq k - 1$ and $\lambda_{\max} \geq k + 1$.

Theorem (Rayleigh's Theorem). Let A be a real symmetric $n \times n$ matrix with all positive eigenvalues, then $\lambda_{\min} \leq \frac{v^\top A v}{v^\top v} \leq \lambda_{\max}$ for every $v \in \mathbb{R}^n, v \neq 0$.

$det(A_2 - \lambda I) = 0 \Rightarrow \lambda_{min} = k-1, \lambda_{max} = k+1$

$(k-\lambda)^3 - 2(k-\lambda) = 0 \Rightarrow \lambda = k$ or $\lambda = k+\sqrt{2}$ or $k - \sqrt{2}$

$det(A_5) = k \cdot det(A_4) - det(A_3)$

$$\frac{v^T A v}{v^T v} = \frac{\sum_{i,j=1\ldots n} a_{ij} v_i v_j}{v_1^2 + v_2^2 + \cdots + v_n^2}$$

Group 1 work

1,2,3,4,6

1. Let λ_1, λ_2 be two different eigenvalues of the symmetric matrix A. And v_1, v_2 are the associated eigenvectors with λ_1, λ_2 respectively. Show that $v_1^T v_2 = 0$.

solution:

A is a symmetric matrix $\Rightarrow A^T = A$.

$$\begin{aligned} Av_1 = \lambda_1 v_1 \Rightarrow \quad & (Av_1)^T &=& \quad (\lambda_1 v_1)^T \\ \Rightarrow \quad & v_1^T A^T &=& \quad \lambda_1 v_1^T \\ \Rightarrow \quad & v_1^T A &=& \quad \lambda_1 v_1^T \\ \Rightarrow \quad & v_1^T A v_2 &=& \quad \lambda_1 v_1^T v_2 \end{aligned}$$

$$Av_2 = \lambda_2 v_2 \Rightarrow \quad v_1^T A v_2 = \lambda_2 v_1^T v_2$$

$$\begin{cases} \lambda_1 - \lambda_2 \neq 0 \\ (\lambda_1 - \lambda_2) v_1^T v_2 = 0 \end{cases} \Rightarrow v_1^T v_2 = 0.$$

2. Suppose A and M are $n \times n$ real matrices with nonzero eigenvalues. A is invertible and $AMA^{-1} = M^2$. Prove that for any nonzero eigenvalues of M, there exists some $m \in \mathbb{Z}_{++}$ such that the eigenvalue is a root of $\alpha^m = 1$. Moreover, there exists some $n \in \mathbb{Z}_{++}$ such that any nonzero eigenvalue of M is the root of $\alpha^n = 1$.

solution:

Suppose λ is a nonzero eigenvalue of M, and $\mathbf{x}$ is an eigenvector corresponding to λ. Then,

$M\mathbf{x} = \lambda \mathbf{x}$

$M^2 \mathbf{x} = M(M\mathbf{x}) = M(\lambda \mathbf{x}) = \lambda(M\mathbf{x}) = \lambda^2 \mathbf{x}$

$\Rightarrow \lambda^2$ is a nonzero eigenvalue of M^2.

$M^2 = AMA^{-1} \Rightarrow M^2$ and M are similar $\Rightarrow$ They have the same eigenvalues.

$\Rightarrow \lambda^2$ is a nonzero eigenvalue of M.

Repeat the process above, we get

$\lambda,\ \lambda^2,\ \lambda^4,\ \cdots,\ \lambda^{2^k} (\lambda \neq 0,\ k \in Z^+,\ k \to \infty)$ are all eigenvalues of M.

Suppose M has t distinct nonzero eigenvalues $\lambda^{2^{p_1}} < \lambda^{2^{p_2}} < \cdots < \lambda^{2^{p_t}}$ in the set $\{\lambda,\ \lambda^2,\ \lambda^4,\ \cdots,\ \lambda^{2^k}\}$. Then,

$\exists\, q \in \mathbb{Z},\ q > p_t$ such that $\lambda^{2^q} = \lambda^{2^p}$, $p \in \{p_1,\ p_2,\ \cdots,\ p_t\}$.

$\lambda^{2^q-2^p}=1.$

$\Rightarrow \exists\, m=2^q-2^p\in\mathbb{Z}_{++}$ such that the eigenvalue is a root of $\alpha^m=1$.

Suppose M has s distinct nonzero eigenvalues $\lambda_1,\ \lambda_2,\ \cdots,\ \lambda_s$ in total.

For each of them, $\exists\, m_1,\ m_2,\ \cdots,\ m_s\in\mathbb{Z}_{++}$, such that

$\lambda_1^{m_1}=1,\ \lambda_2^{m_2}=1,\ \cdots,\ \lambda_s^{m_s}=1.$

$\Rightarrow \exists\, n=m_1m_2\cdots m_s\in\mathbb{Z}_{++}$ such that any nonzero eigenvalue of M is the root of $\alpha^n=1$.

3. You can use the following observation without proof: For any linear transformation T, if there exists a T-invariant subspace V, then T must have an eigenvector in V.

Let A and B are real $n\times n$ matrices such that both are with real eigenvalues; moreover, $AB=BA$. Prove that A and B have a common eigenvector.

solution:

Suppose S_λ is an eigenspace of A corresponding to the eigenvalue λ.

$$\begin{aligned}\forall\alpha\in S_\lambda,\ A\alpha&=\lambda\alpha\\ BA\alpha&=B\lambda\alpha\\ A(B\alpha)&=\lambda(B\alpha)\end{aligned}$$

$\Rightarrow B\alpha\in S_\lambda.$

S_λ is B-invariant $\Rightarrow B$ must have an eigenvector in S_λ.

$\Rightarrow A$ and B have a common eigenvector.

4. Let V be the vector space of sequences (a_n) of real numbers. The shift operator $S: V\to V$ is defined by

$$S\left((a_1,a_2,a_3,\ldots)\right)=(a_2,a_3,a_4,\ldots).$$

(a) Given an eigenvalue of S being λ, find the corresponding eigenvectors.

(b) Show that the subspace W consisting of the sequences (x_n) with $x_{n+2}=x_{n+1}+x_n$ is a twodimensional, S-invariant subspace of V and exhibit an explicit basis for W.

(c) Find an explicit formula for the Fibonacci number f_n, where $f_2=f_1=1, f_{n+2}=f_{n+1}+f_n$ for $n\geq 1$.

solution:

(a) Suppose $\mathbf{v}$ is an eigenvector corresponding to λ. Then,

$$\lambda\mathbf{v}=S(\mathbf{v})$$

$$\begin{bmatrix}\lambda v_1\\ \lambda v_2\\ \lambda v_3\\ \vdots\end{bmatrix}=\begin{bmatrix}v_2\\ v_3\\ v_4\\ \vdots\end{bmatrix}$$

$$\Rightarrow \mathbf{v}=\begin{bmatrix}v_1\\ \lambda v_1\\ \lambda^2 v_1\\ \lambda^3 v_1\\ \vdots\end{bmatrix}=v_1\begin{bmatrix}1\\ \lambda\\ \lambda^2\\ \lambda^3\\ \vdots\end{bmatrix},\ v_1\in\mathbb{R},\ v_1\neq 0.$$

(b) $W=\{(x_1,x_2,\cdots)\in V|x_{n+2}=x_{n+1}+x_n\}$, whose elements are only determined by the first two entries.

$\Rightarrow dim(W)=2$ and clearly $\{(0,1,\cdots),(1,0,\cdots)\}$ is a basis for W.

assume in W, there exist two eigenvectors of S, $\mathbf{p}$ and $\mathbf{q}$, corresponding to distinct eigenvalues $\lambda_p,\ \lambda_q$.

Let $\mathbf{p} = \begin{bmatrix} 1 \\ \lambda_p \\ \lambda_p^2 \\ \vdots \end{bmatrix}$, $\mathbf{q} = \begin{bmatrix} 1 \\ \lambda_q \\ \lambda_q^2 \\ \vdots \end{bmatrix}$.

$x^2 = x + 1 \Rightarrow \lambda_p = \frac{1+\sqrt{5}}{2},\ \lambda_q = \frac{1-\sqrt{5}}{2}$.

The assumption is correct.

$\mathbf{p},\ \mathbf{q}$ are linearly independent $\Rightarrow \{\mathbf{p}, \mathbf{q}\}$ is also a basis for W.

$\forall\, \mathbf{x} \in W, \exists\, \alpha, \beta \in \mathbb{R}$ such that

$\mathbf{x} = \alpha\mathbf{p} + \beta\mathbf{q}$.

$\forall \lambda \in \mathbb{R},\ \mathbf{u}, \mathbf{v} \in V,$

$\begin{cases} S(\lambda\mathbf{v}) = \lambda S(\mathbf{v}) \\ S(\mathbf{u}+\mathbf{v}) = S(\mathbf{u}) + S(\mathbf{v}) \end{cases} \Rightarrow S$ is a linear mapping. Then,

$$\begin{aligned} S(x) &= \alpha S(\mathbf{p}) + \beta S(\mathbf{q}) \\ &= \alpha\lambda_p\mathbf{p} + \beta\lambda_q\mathbf{q} \in W \end{aligned}$$

$\forall\, \mathbf{x} \in W \Rightarrow S(\mathbf{x}) \in W \Rightarrow W$ is S-invariant.

(c)

$$\mathbf{f} = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 3 \\ \vdots \end{bmatrix} = \alpha\mathbf{p} + \beta\mathbf{q}$$

$$\begin{cases} \alpha + \beta = 1 \\ \lambda_p\alpha + \lambda_q\beta = 1 \end{cases} \Rightarrow \begin{cases} \alpha = \dfrac{1}{\sqrt{5}} \cdot \dfrac{\sqrt{5}+1}{2} \\ \beta = \dfrac{1}{\sqrt{5}} \cdot \dfrac{\sqrt{5}-1}{2} \end{cases}$$

$$\begin{aligned} f_n &= \alpha p_n + \beta q_n \\ &= \alpha\lambda_p^{n-1} + \beta\lambda_q^{n-1} \\ &= \frac{1}{\sqrt{5}} \cdot \frac{\sqrt{5}+1}{2}(\frac{1+\sqrt{5}}{2})^{n-1} + \frac{1}{\sqrt{5}} \cdot \frac{\sqrt{5}-1}{2}(\frac{1-\sqrt{5}}{2})^{n-1} \\ &= \frac{1}{\sqrt{5}}[(\frac{1+\sqrt{5}}{2})^n - (\frac{1-\sqrt{5}}{2})^n] \end{aligned}$$

6. Let A be a $n \times n$ matrix such that: 1) all eigenvalues are positive; 2) A is diagonalizable with orthogonal matrix. Let $y \in \mathbb{R}^n, y \neq \mathbf{0}$. Prove that the limit

$$\lim_{m\to\infty} \frac{y^\top A^{m+1} y}{y^\top A^m y}$$

exists and is an eigenvalue of A.

solution:

A is diagonalizable with orthogonal matrix. $\Rightarrow A = P^{-1}BP,\ B = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}$, $P^{-1} = P^T$.

$A^k = P^T B^k P$

$y^T A^k y = y^T P^T B^k P y = (Py)^T B^k (Py)$

P is inversible and $y \neq 0 \Rightarrow Py \neq 0$. Suppose $Py = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$.

$$
\begin{aligned}
(Py)^T B^k (Py) &= [\, y_1 \quad y_2 \quad \cdots \quad y_n \,] \begin{bmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n^k \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \\
&= y_1^2 \lambda_1^k + y_2^2 \lambda_2^k + \cdots + y_n^2 \lambda_n^k
\end{aligned}
$$

$$
\begin{aligned}
\lim_{m \to \infty} \frac{y^\top A^{m+1} y}{y^\top A^m y} &= \lim_{m \to \infty} \frac{y_1^2 \lambda_1^{m+1} + y_2^2 \lambda_2^{m+1} + \cdots + y_n^2 \lambda_n^{m+1}}{y_1^2 \lambda_1^m + y_2^2 \lambda_2^m + \cdots + y_n^2 \lambda_n^m} \\
&= \lambda_{max} \lim_{m \to \infty} \frac{y_1^2 (\frac{\lambda_1}{\lambda_{max}})^{m+1} + y_2^2 (\frac{\lambda_2}{\lambda_{max}})^{m+1} + \cdots + y_n^2 (\frac{\lambda_n}{\lambda_{max}})^{m+1}}{y_1^2 (\frac{\lambda_1}{\lambda_{max}})^m + y_2^2 (\frac{\lambda_2}{\lambda_{max}})^m + \cdots + y_n^2 (\frac{\lambda_n}{\lambda_{max}})^m} \\
&= \lambda_{max}
\end{aligned}
$$

exercise 9

1. (a) Let W be a subspace of $\mathbb{R}^m$, and let $V = \{v_1, v_2, \ldots, v_n\}$ be an orthogonal subset of $W, m \geq n$. Prove that for any $x \in \text{Span}(V)$ we have

$$\|x\|^2 = \sum_{i=1}^{n} \frac{(x \cdot v_i)^2}{\|v_i\|^2}$$

(b) (Bessel's Inequality) In the context of (a), prove that for any $x \in W$ we have

$$\|x\|^2 \geq \sum_{i=1}^{n} \frac{(x \cdot v_i)^2}{\|v_i\|^2}$$

2. Let W_1 and W_2 be subspaces $\mathbb{R}^n$. Recalling that $W_1 + W_2 = \{u + v : u \in W_1, v \in W_2\}$, prove that $(W_1 + W_2)^\perp = W_1^\perp \cap W_2^\perp$ and $(W_1 \cap W_2)^\perp = W_1^\perp + W_2^\perp$.

3. Let A be an $m \times n$ matrix. Prove that every vector x in $\mathbb{R}^n$ can be written in the form $x = p + u$, where $p \in \text{Row}(A), u \in \text{Null}(A)$. Also, show that if the equation $Ax = b$ is consistent, then there is a unique p in $\text{Row}(A)$ such that $Ap = b$.

4. Suppose $A = QR$, where Q is an $m \times n$ matrix and R is an $n \times n$ matrix.

(a) Show that if the columns of A are linearly independent, then R must be invertible.

(b) Show that if R is invertible, then A and Q have the same column space.

5. Suppose $A = QR$ is a QR factorization of an $m \times n$ matrix A (with linearly independent columns). Partition A as $[A_1 A_2]$, where A_1 has p columns. Show how to obtain a QR factorization of A_1, and explain why your factorization has the appropriate properties.

6. In each of the following parts, find an orthogonal basis of the given subspace of W and calculate the orthogonal projection of the given element $u \in V$ on W.

(a) $V = \mathbb{R}^3, u = (2,1,3)$, and $W = \{(x,y,z) : x + 3y - 2z = 0\}$.

(b) $V = P(\mathbb{R})$, i.e., the space of polynomial functions in $\mathbb{R}$. Let V be with the inner product $f \cdot g = \int_0^1 f(t)g(t)dt, u(x) = 4 + 3x - 2x^2$, and $W = P_1(\mathbb{R})$, i.e., the space of linear functions in $\mathbb{R}$.

7. Let $V = C([0,1])$ be the space of continuous functions on $[0,1]$. Let V be with the inner product $f \cdot g = \int_0^1 f(t)g(t)dt$. Let W be the subspace spanned by the linearly independent set $\{t, \sqrt{t}\}$.

(a) Use the Gram-Schmidt process to find an orthonormal basis for W.

(b) Let $h(t) = t^2$. Use the orthonormal basis obtained in (a) to obtain the "best" (closest) approximation of h in W.

8. For continuous real valued functions f, g on the interval $[-1,1]$, define the inner product $f \cdot g = \int_{-1}^1 f(x)g(x)dx$. Find that polynomial of the form $p(x) = a + bx^2 - x^4$ which is orthogonal on $[-1,1]$ to all polynomials with order $0, 1, 2, 3$

Group1: 12345

Group 1 work

1. (a) Let W be a subspace of $\mathbb{R}^m$, and let $V = \{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$ be an orthogonal subset of $W, m \geq n$. Prove that for any $\mathbf{x} \in \mathrm{Span}(V)$ we have

$$\|\mathbf{x}\|^2 = \sum_{i=1}^{n} \frac{(\mathbf{x} \cdot \mathbf{v_i})^2}{\|\mathbf{v_i}\|^2}$$

(b) (Bessel's Inequality) In the context of (a), prove that for any $\mathbf{x} \in W$ we have

$$\|\mathbf{x}\|^2 \geq \sum_{i=1}^{n} \frac{(\mathbf{x} \cdot \mathbf{v_i})^2}{\|\mathbf{v_i}\|^2}$$

solution:

(a)

$\forall\, \mathbf{x} \in Span(V)$, $\mathbf{x} = c_1\mathbf{v_1} + c_2\mathbf{v_2} + \dots + c_n\mathbf{v_n}$.

$$\begin{aligned}\Rightarrow \|\mathbf{x}\|^2 &= \mathbf{x} \cdot \mathbf{x} \\ &= (c_1\mathbf{v_1} + c_2\mathbf{v_2} + \dots + c_n\mathbf{v_n}) \cdot (c_1\mathbf{v_1} + c_2\mathbf{v_2} + \dots + c_n\mathbf{v_n}) \\ &= c_1^2\|\mathbf{v_1}\|^2 + c_2^2\|\mathbf{v_2}\|^2 + \dots + c_n^2\|\mathbf{v_n}\|^2 \\ &= \sum_{i=1}^{n} c_i^2\|\mathbf{v_i}\|^2\end{aligned}$$

$$\begin{aligned}\sum_{i=1}^{n} \frac{(\mathbf{x} \cdot \mathbf{v_i})^2}{\|\mathbf{v_i}\|^2} &= \sum_{i=1}^{n} \frac{\left(c_i\|\mathbf{v_i}\|^2\right)^2}{\|\mathbf{v_i}\|^2} \\ &= \sum_{i=1}^{n} c_i^2\|\mathbf{v_i}\|^2\end{aligned}$$

Hence, $\|\mathbf{x}\|^2 = \sum_{i=1}^{n} \frac{(\mathbf{x} \cdot \mathbf{v_i})^2}{\|\mathbf{v_i}\|^2}$

(b)

for any $\mathbf{x} \in W$, it can be written uniquely in the form

$\mathbf{x} = \hat{\mathbf{x}} + \mathbf{z}$,

where $\hat{\mathbf{x}} \in Span(V)$ and $\mathbf{z} \in Span(V)^{\perp}$.

Because $\|\mathbf{z}\| \geq 0$, by part (a) we have

$$\begin{aligned}\|\mathbf{x}\|^2 &= \|\hat{\mathbf{x}}\|^2 + \|\mathbf{z}\|^2 \\ &\geq \sum_{i=1}^{n} \frac{(\mathbf{x} \cdot \mathbf{v_i})^2}{\|\mathbf{v_i}\|^2}\end{aligned}$$

2. Let W_1 and W_2 be subspaces $\mathbb{R}^n$. Recalling that $W_1 + W_2 = \{\mathbf{u} + \mathbf{v} : \mathbf{u} \in W_1, \mathbf{v} \in W_2\}$, prove that $(W_1 + W_2)^{\perp} = W_1^{\perp} \cap W_2^{\perp}$ and $(W_1 \cap W_2)^{\perp} = W_1^{\perp} + W_2^{\perp}$.

solution:

$\forall\, \mathbf{x} \in (W_1 + W_2)^{\perp}$, we have

$\mathbf{x} \cdot (\mathbf{u} + \mathbf{v}) = 0$ for any $\mathbf{u} \in W_1$ and $\mathbf{v} \in W_2$.

Let $\mathbf{u} = \mathbf{0} \Rightarrow \mathbf{x} \cdot \mathbf{v} = 0$ for any $\mathbf{v} \in W_2 \Rightarrow \mathbf{x} \in W_2^{\perp}$.

Let $\mathbf{v} = \mathbf{0} \Rightarrow \mathbf{x} \cdot \mathbf{u} = 0$ for any $\mathbf{u} \in W_1 \Rightarrow \mathbf{x} \in W_1^{\perp}$.

Hence $\mathbf{x} \in W_1^{\perp} \cap W_2^{\perp} \Rightarrow (W_1 + W_2)^{\perp} \subseteq W_1^{\perp} \cap W_2^{\perp}$.

$\forall\, \mathbf{x} \in W_1^{\perp} \cap W_2^{\perp} \Rightarrow \mathbf{x} \in W_1^{\perp}$ and $\mathbf{x} \in W_2^{\perp}$.

$\mathbf{x} \cdot (\mathbf{u} + \mathbf{v}) = \mathbf{x} \cdot \mathbf{u} + \mathbf{x} \cdot \mathbf{v} = 0 + 0 = 0$ for any $\mathbf{u} \in W_1$ and $\mathbf{v} \in W_2$.

Hence $\mathbf{x} \in (W_1 + W_2)^{\perp} \Rightarrow W_1^{\perp} \cap W_2^{\perp} \subseteq (W_1 + W_2)^{\perp}$.

$$\begin{cases} (W_1 + W_2)^{\perp} \subseteq W_1^{\perp} \cap W_2^{\perp} \\ W_1^{\perp} \cap W_2^{\perp} \subseteq (W_1 + W_2)^{\perp} \end{cases} \Rightarrow (W_1 + W_2)^{\perp} = W_1^{\perp} \cap W_2^{\perp}.$$

Suppose W is a subspace of $\mathbb{R}^n$.

$\forall\, \mathbf{x} \in W,\ \mathbf{y} \in W^{\perp}$, we have

$\mathbf{x} \cdot \mathbf{y} = 0 \Rightarrow \mathbf{x} \in (W^{\perp})^{\perp} \Rightarrow W \subseteq (W^{\perp})^{\perp}$.

$\forall\, \mathbf{x} \in (W^{\perp})^{\perp}$, it can be written uniquely in the form

$\mathbf{x} = \hat{\mathbf{x}} + \mathbf{z}$, where $\hat{\mathbf{x}} \in W$ and $\mathbf{z} \in W^{\perp}$. Because $\hat{\mathbf{x}} \cdot \mathbf{z} = 0$, we have

$\mathbf{x} \cdot \mathbf{z} = (\hat{\mathbf{x}} + \mathbf{z}) \cdot \mathbf{z} = \mathbf{z}^2 = 0 \Rightarrow \mathbf{z} = \mathbf{0}$.

$\mathbf{x} = \hat{\mathbf{x}} \in W \Rightarrow (W^{\perp})^{\perp} \subseteq W$.

$$\begin{cases} W \subseteq (W^{\perp})^{\perp} \\ (W^{\perp})^{\perp} \subseteq W \end{cases} \Rightarrow (W^{\perp})^{\perp} = W.$$

$$\begin{aligned} (W_1 + W_2)^{\perp} &= W_1^{\perp} \cap W_2^{\perp} \\ \Rightarrow \left(W_1^{\perp} + W_2^{\perp}\right)^{\perp} &= (W_1^{\perp})^{\perp} \cap (W_2^{\perp})^{\perp} \\ \Rightarrow \left(W_1^{\perp} + W_2^{\perp}\right)^{\perp} &= W_1 \cap W_2 \\ \Rightarrow (\left(W_1^{\perp} + W_2^{\perp}\right)^{\perp})^{\perp} &= (W_1 \cap W_2)^{\perp} \\ \Rightarrow W_1^{\perp} + W_2^{\perp} &= (W_1 \cap W_2)^{\perp} \end{aligned}$$

3. Let A be an $m \times n$ matrix. Prove that every vector $\mathbf{x}$ in $\mathbb{R}^n$ can be written in the form $\mathbf{x} = \mathbf{p} + \mathbf{u}$, where $\mathbf{p} \in \text{Row}(A), \mathbf{u} \in \text{Null}(A)$. Also, show that if the equation $A\mathbf{x} = \mathbf{b}$ is consistent, then there is a unique $\mathbf{p}$ in $\text{Row}(A)$ such that $A\mathbf{p} = \mathbf{b}$.

solution:

$(Row(A))^{\perp} = Null(A)$

$\Rightarrow$ every vector $\mathbf{x}$ in $\mathbb{R}^n$ can be written in the form $\mathbf{x} = \mathbf{p} + \mathbf{u}$, where $\mathbf{p} \in \text{Row}(A), \mathbf{u} \in \text{Null}(A)$.

If the equation $A\mathbf{x} = \mathbf{b}$ is consistent, suppose $\mathbf{y}$ is a solution of the equation.

$\mathbf{y}$ can be written in the form $\mathbf{y} = \mathbf{p} + \mathbf{u}$, where $\mathbf{p} \in \text{Row}(A), \mathbf{u} \in \text{Null}(A)$.

$A\mathbf{p} = A(\mathbf{y} - \mathbf{u}) = A\mathbf{y} - A\mathbf{u} = A\mathbf{y} = \mathbf{b}$

$\Rightarrow \exists\, \mathbf{p} \in Row(A)$ such that $A\mathbf{p} = \mathbf{b}$.

Suppose there exists $\mathbf{p\prime} \in Row(A)$ such that $A\mathbf{p\prime} = \mathbf{b}$, then

$A(\mathbf{p} - \mathbf{p\prime}) = \mathbf{0} \Rightarrow \mathbf{p} - \mathbf{p\prime} \in Null(A)$.

$\begin{cases} \mathbf{p} - \mathbf{p\prime} \in Row(A) \\ \mathbf{p} - \mathbf{p\prime} \in Null(A) \end{cases} \Rightarrow \mathbf{p} - \mathbf{p\prime} \in Row(A) \cap Null(A)$.

$(Row(A))^{\perp} = Null(A) \Rightarrow (\mathbf{p} - \mathbf{p\prime}) \cdot (\mathbf{p} - \mathbf{p\prime}) = 0 \Rightarrow \mathbf{p} - \mathbf{p\prime} = \mathbf{0}$.

Hence $\mathbf{p\prime} = \mathbf{p} \Rightarrow$ there is a unique $\mathbf{p}$ in $\text{Row}(A)$ such that $A\mathbf{p} = \mathbf{b}$.

4. Suppose $A = QR$, where Q is an $m \times n$ matrix and R is an $n \times n$ matrix.

(a) Show that if the columns of A are linearly independent, then R must be invertible.

(b) Show that if R is invertible, then A and Q have the same column space.

solution:

(a)

If $R\mathbf{x} = \mathbf{0}$, then

$A\mathbf{x} = QR\mathbf{x} = Q(R\mathbf{x}) = \mathbf{0}$.

the columns of A are linearly independent $\Rightarrow \mathbf{x} = \mathbf{0}$.

Then R must be invertible.

(b) If R is invertible,

$\forall\, \mathbf{y} \in Col(A)$, we have

$\mathbf{y} = A\mathbf{x}$, where $\mathbf{x} \in \mathbb{R}^n$.

$\mathbf{y} = A\mathbf{x} = QR\mathbf{x} = Q(R\mathbf{x})$

$\Rightarrow \mathbf{y} \in Col(Q)$.

Hence $Col(A) \subseteq Col(Q)$.

$\forall\, \mathbf{y} \in Col(Q)$, we have

$\mathbf{y} = Q\mathbf{x}$, where $\mathbf{x} \in \mathbb{R}^n$.

$\mathbf{y} = Q\mathbf{x} = QRR^{-1}\mathbf{x} = A(R^{-1}\mathbf{x})$

$\Rightarrow \mathbf{y} \in Col(A)$.

Hence $Col(Q) \subseteq Col(A)$.

$$\begin{cases} Col(A) \subseteq Col(Q) \\ Col(Q) \subseteq Col(A) \end{cases} \Rightarrow Col(Q) = Col(A).$$

5. Suppose $A = QR$ is a QR factorization of an $m \times n$ matrix A (with linearly independent columns). Partition A as $[A_1 A_2]$, where A_1 has p columns. Show how to obtain a QR factorization of A_1, and explain why your factorization has the appropriate properties.

solution:

To obtain a QR factorization of A_1, we can partition Q and R like this:

$$A = QR = [\, Q_1 \quad Q_2 \,] \begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{bmatrix},$$

where Q_1 is $m \times p$, Q_2 is $m \times (n-p)$, R_{11} is $p \times p$, R_{12} is $p \times (n-p)$, R_{22} is $(n-p) \times (n-p)$.

R_{11} and R_{22} are upper triangular invertible matrix with positive entries on its diagonal.

$$A = [\, Q_1 \quad Q_2 \,] \begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{bmatrix} = [\, Q_1 R_{11} \quad Q_1 R_{12} + Q_2 R_{22} \,].$$

Then $A_1 = Q_1 R_{11}$.

Because the columns of Q form an orthonormal basis, the first p columns of Q also form an an orthonormal basis.

In addition, R_{11} is an upper triangular invertible matrix with positive entries on its diagonal.

Hence this factorization has the appropriate properties.

exercise 10

1. Let $(a, b, c) \in \mathbb{R}^3$ be a vector of length 1 . Let W be the plane defined by $ax + by + cz = 0$. Find, in the standard basis, the matrix representing the orthogonal projection of $\mathbb{R}^3$ onto W.

2. Let w be a positive continuous function on $[0, 1]$, n be a positive integer, and P_n be the vector space of real polynomials whose degrees are at most n. Define the inner product

$$\langle p, q \rangle = \int_0^1 p(t)q(t)w(t)dt$$

for all $p, q \in P_n$.

(a) Prove that P_n has an orthonormal basis $p_0, p_1, \ldots, p_n$ (i.e., $\langle p_j, p_k \rangle = 1$ for $j = k$ and 0 for $j \neq k$) such that the degree of p_k is k, for each k.

(b) Let p_k' be the derivative of p_k. Prove that $\langle p_k, p_k' \rangle = 0$ for each k.

3. Let a and b be real numbers. Prove that there are two orthogonal unit vectors u and v in $\mathbb{R}^3$ such that $u = (u_1, u_2, a)$ and $v = (v_1, v_2, b)$ if and only if $a^2 + b^2 \leq 1$.

4. In physics, Hooke's law states that (within certain limits) there is a linear relationship between the length x of a spring and the force y applied to (or exerted by) the spring. That is, $y = cx + d$, where c is called the spring constant. Use the following data to estimate the spring constant (the length is given in inches and the force is given in pounds).

\begin{tabular}{cc}
\hline Length x & Force y \
\hline 3.5 & 1.0 \
4.0 & 2.2 \
4.5 & 2.8 \
5.0 & 4.3 \
\hline
\end{tabular}

5. For the data $\{(-2,4),(-1,3),(0,1),(1,-1),(2,-3)\}$, use the least squares approximation to find the best fits with both (i) a linear function and (ii) a quadratic function. Compute the error E in both cases.

6. For f, g in $C[a, b]$, set

$$\langle f, g\rangle = \int_a^b f(t)g(t)dt$$

Let V be the space $C[-1,1]$ with the inner product defined above. Find an orthogonal basis for the subspace spanned by the polynomials $1, t$ and t^2. The polynomials in this basis are called Legendre polynomials.

7. Prove the following facts related to positive definite and eigenvalue.

(a) If B is $m \times n$, then $B^\top B$ is positive semidefinite; if B is $n \times n$ and invertible, then $B^\top B$ is positive definite.

(b) If an $n \times n$ matrix A is positive definite, then there exists a positive definite matrix B such that $A = B^\top B$. [Hint: Write $A = PDP^\top$, with $P^\top = P^{-1}$.]

(c) Let A and B be symmetric $n \times n$ matrices whose eigenvalues are all positive. Show that the eigenvalues of $A + B$ are all positive.

(d) Let A be an $n \times n$ invertible symmetric matrix. Show that if A is positive definite, then so is A^{-1}. [Hint: Consider eigenvalues.]

Group 1 12345

Group 1 work

1. Let $(a, b, c) \in \mathbb{R}^3$ be a vector of length 1 . Let W be the plane defined by $ax + by + cz = 0$. Find, in the standard basis, the matrix representing the orthogonal projection of $\mathbb{R}^3$ onto W.

   solution:

   Define P_W as the matrix representing the orthogonal projection of $\mathbb{R}^3$ onto W.

   Define $P_{W^\perp}$ as the matrix representing the orthogonal projection of $\mathbb{R}^3$ onto $W^\perp$.

   $\forall\ \mathbf{y} \in \mathbb{R}^3$, we have

   $$\begin{aligned} \mathbf{y} &= proj_W \mathbf{y} + proj_{W^\perp}\mathbf{y} \\ &= P_W\mathbf{y} + P_{W^\perp}\mathbf{y} \\ &= (P_W + P_{W^\perp})\mathbf{y} \end{aligned}$$
   $$\Rightarrow P_W + P_{W^\perp} = I.$$

   $$P_{W^\perp}\mathbf{e_1} = \frac{\mathbf{e_1}\cdot\begin{bmatrix} a \\ b \\ c \end{bmatrix}}{\begin{bmatrix} a \\ b \\ c \end{bmatrix}\cdot\begin{bmatrix} a \\ b \\ c \end{bmatrix}}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = a\begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

Similarly,

$$P_{W^\perp}\mathbf{e_2} = b\begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

$$P_{W^\perp}\mathbf{e_3} = c\begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

$$P_{W^\perp} = \begin{bmatrix} a\begin{bmatrix} a \\ b \\ c \end{bmatrix} & b\begin{bmatrix} a \\ b \\ c \end{bmatrix} & c\begin{bmatrix} a \\ b \\ c \end{bmatrix} \end{bmatrix}$$

$$= \begin{bmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{bmatrix}$$

$$P_W = I - P_{W^\perp} = \begin{bmatrix} 1-a^2 & -ab & -ac \\ -ab & 1-b^2 & -bc \\ -ac & -bc & 1-c^2 \end{bmatrix}.$$

2. Let w be a positive continuous function on $[0,1]$, n be a positive integer, and P_n be the vector space of real polynomials whose degrees are at most n. Define the inner product

$$\langle p, q\rangle = \int_0^1 p(t)q(t)w(t)dt$$

for all $p, q \in P_n$.

(a) Prove that P_n has an orthonormal basis $p_0, p_1, \ldots, p_n$ (i.e., $\langle p_j, p_k\rangle = 1$ for $j = k$ and 0 for $j \neq k$) such that the degree of p_k is k, for each k.

(b) Let p'_k be the derivative of p_k. Prove that $\langle p_k, p'_k\rangle = 0$ for each k.

solution:

(a)

$c_0 + c_1t + c_2t^2 + \cdots + c_nt^n = 0$ if and only if $c_0 = c_1 = \cdots = c_n = 0$.

Also, $\forall\ p \in P_n$, we have $p = c_0 + c_1t + \cdots + c_nt^n$.

$\Rightarrow \{1,\ t,\ t^2,\ \cdots,\ t^n\}$ is a basis of P_n.

By the Gram-Schmidt process,

$v_0 = 1$, whose degree is 0.

$v_1 = t - \frac{t\cdot v_0}{v_0\cdot v_0}v_0$, whose degree is 1.

$\vdots$

$v_n = t^n - \frac{t^n\cdot v_0}{v_0\cdot v_0}v_0 - \frac{t^n\cdot v_1}{v_1\cdot v_1}v_1 - \cdots - \frac{t^n\cdot v_{n-1}}{v_{n-1}\cdot v_{n-1}}v_{n-1}$, whose degree is n.

$\{v_0,\ v_1,\ \cdots,\ v_n\}$ is a orthogonal basis of P_n.

Normalize it, we get the orthonormal basis $\{p_0,\ p_1,\ \cdots,\ p_n\}$ where

$p_0 = \frac{v_0}{||v_0||}$, whose degree is 0.

$p_1 = \frac{v_1}{||v_1||}$, whose degree is 1.

$\vdots$

$p_n = \frac{v_n}{||v_n||}$, whose degree is n.

(b)

$p'_k = \frac{d}{dt}p_k$, whose degree is $k-1$.

$\Rightarrow p'_k = c_0 + c_1t + \cdots + c_{k-1}t^{k-1} \Rightarrow p'_k \in Span\{1,\ t,\ \cdots,\ t^{k-1}\}$.

By the Gram-Schmidt process, we know

the orthonormal set $\{p_0,\ p_1,\ \cdots,\ p_{k-1}\}$ is also a basis of $Span\{1,\ t,\ \cdots,\ t^{k-1}\}$.

$\Rightarrow p'_k = c'_0 p_0 + c'_1 p_1 + \cdots + c'_{k-1} p_{k-1}$.

$\langle p_k, p'_k \rangle = p_k \cdot (c'_0 p_0 + c'_1 p_1 + \cdots + c'_{k-1} p_{k-1}) = 0$.

3. Let a and b be real numbers. Prove that there are two orthogonal unit vectors u and v in $\mathbb{R}^3$ such that $u = (u_1, u_2, a)$ and $v = (v_1, v_2, b)$ if and only if $a^2 + b^2 \leq 1$.

solution:

(i) Prove there are two orthogonal unit vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^3$ such that $\mathbf{u} = (u_1, u_2, a)$ and $\mathbf{v} = (v_1, v_2, b) \Rightarrow a^2 + b^2 \leq 1$.

Suppose $W = Span\{\mathbf{u},\ \mathbf{v}\}, \mathbf{e_3} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

$proj_W \mathbf{e_3} = \frac{\mathbf{e_3} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}} \mathbf{u} + \frac{\mathbf{e_3} \cdot \mathbf{v}}{\mathbf{v} \cdot \mathbf{v}} \mathbf{v} = a\mathbf{u} + b\mathbf{v}$.

$||proj_W \mathbf{e_3}|| \leq ||\mathbf{e_3}|| \Rightarrow a^2 + b^2 \leq 1$.

(ii) Prove $a^2 + b^2 \leq 1 \Rightarrow$ there are two orthogonal unit vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^3$ such that $\mathbf{u} = (u_1, u_2, a)$ and $\mathbf{v} = (v_1, v_2, b)$.

When $a^2 + b^2 \leq 1$, there exists a real number c such that $a^2 + b^2 + c^2 = 1$.

Let $\mathbf{u} = \begin{bmatrix} \sqrt{b^2 + c^2} \\ 0 \\ a \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} \frac{-ab}{\sqrt{b^2+c^2}} \\ \frac{c}{\sqrt{b^2+c^2}} \\ b \end{bmatrix}$, $\mathbf{w} = \begin{bmatrix} \frac{-ac}{\sqrt{b^2+c^2}} \\ \frac{-b}{\sqrt{b^2+c^2}} \\ c \end{bmatrix}$, then

$\begin{cases} ||\mathbf{u}|| = ||\mathbf{v}|| = ||\mathbf{w}|| = 1 \\ \mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} = 0 \end{cases} \Rightarrow \{\mathbf{u},\ \mathbf{v},\ \mathbf{w}\}$ is a orthonormal basis of $\mathbb{R}^3$.

Then, there exist two orthogonal unit vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^3$, whose third entries are a and b respectively.

4. In physics, Hooke's law states that (within certain limits) there is a linear relationship between the length x of a spring and the force y applied to (or exerted by) the spring. That is, $y = cx + d$, where c is called the spring constant. Use the following data to estimate the spring constant (the length is given in inches and the force is given in pounds).

Length (x)	Force (y)
3.5	1.0
4.0	2.2
4.5	2.8
5.0	4.3

Suppose $A = \begin{bmatrix} 3.5 & 1 \\ 4.0 & 1 \\ 4.5 & 1 \\ 5.0 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 1.0 \\ 2.2 \\ 2.8 \\ 4.3 \end{bmatrix}$, $\mathbf{m} = \begin{bmatrix} c \\ d \end{bmatrix}$.

Then the least-squares solution of $A\mathbf{m} = \mathbf{b}$ gives the estimation of the the spring constant c.

$$
\begin{aligned}
\hat{\mathbf{m}} &= (A^TA)^{-1}A^T\mathbf{b} \\
&= (\begin{bmatrix} 3.5 & 4.0 & 4.5 & 5.0 \\ 1 & 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 3.5 & 1 \\ 4.0 & 1 \\ 4.5 & 1 \\ 5.0 & 1 \end{bmatrix})^{-1}\begin{bmatrix} 3.5 & 4.0 & 4.5 & 5.0 \\ 1 & 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1.0 \\ 2.2 \\ 2.8 \\ 4.3 \end{bmatrix} \\
&= \begin{bmatrix} 73.5 & 17 \\ 17 & 4 \end{bmatrix}^{-1}\begin{bmatrix} 3.5 & 4.0 & 4.5 & 5.0 \\ 1 & 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1.0 \\ 2.2 \\ 2.8 \\ 4.3 \end{bmatrix} \\
&= \begin{bmatrix} 0.8 & -3.4 \\ -3.4 & 14.7 \end{bmatrix}\begin{bmatrix} 3.5 & 4.0 & 4.5 & 5.0 \\ 1 & 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1.0 \\ 2.2 \\ 2.8 \\ 4.3 \end{bmatrix} \\
&= \begin{bmatrix} -0.6 & -0.2 & 0.2 & 0.6 \\ 2.8 & 1.1 & -0.6 & -2.3 \end{bmatrix}\begin{bmatrix} 1.0 \\ 2.2 \\ 2.8 \\ 4.3 \end{bmatrix} \\
&= \begin{bmatrix} 2.1 \\ -6.35 \end{bmatrix}
\end{aligned}
$$

The estimation of the spring constant c is 2.1 pounds/inch.

5. For the data $\{(-2,4),(-1,3),(0,1),(1,-1),(2,-3)\}$, use the least squares approximation to find the best fits with both (i) a linear function and (ii) a quadratic function. Compute the error E in both cases.

solution:

(i)

Suppose the linear function $f(x) = m_1x + m_2$.

Suppose $A = \begin{bmatrix} -2 & 1 \\ -1 & 1 \\ 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 4 \\ 3 \\ 1 \\ -1 \\ -3 \end{bmatrix}$, $\mathbf{m} = \begin{bmatrix} m_1 \\ m_2 \end{bmatrix}$.

Then the least-squares solution of $A\mathbf{m} = \mathbf{b}$ gives the best fits.

$$\hat{\mathbf{m}} = (A^TA)^{-1}A^T\mathbf{b}$$

$$= \left(\begin{bmatrix} -2 & -1 & 0 & 1 & 2 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} -2 & 1 \\ -1 & 1 \\ 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix}\right)^{-1}\begin{bmatrix} -2 & -1 & 0 & 1 & 2 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 4 \\ 3 \\ 1 \\ -1 \\ -3 \end{bmatrix}$$

$$= \begin{bmatrix} 10 & 0 \\ 0 & 5 \end{bmatrix}^{-1}\begin{bmatrix} -2 & -1 & 0 & 1 & 2 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 4 \\ 3 \\ 1 \\ -1 \\ -3 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{10} & 0 \\ 0 & \frac{1}{5} \end{bmatrix}\begin{bmatrix} -2 & -1 & 0 & 1 & 2 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 4 \\ 3 \\ 1 \\ -1 \\ -3 \end{bmatrix}$$

$$= \begin{bmatrix} -\frac{1}{5} & -\frac{1}{10} & 0 & \frac{1}{10} & \frac{1}{5} \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \end{bmatrix}\begin{bmatrix} 4 \\ 3 \\ 1 \\ -1 \\ -3 \end{bmatrix}$$

$$= \begin{bmatrix} -\frac{9}{5} \\ \frac{4}{5} \end{bmatrix}$$

$f(x) = -\frac{9}{5}x + \frac{4}{5}$.

$$E = ||\mathbf{b} - A\hat{\mathbf{m}}||$$

$$= \left|\left|\begin{bmatrix} 4 \\ 3 \\ 1 \\ -1 \\ -3 \end{bmatrix} - \begin{bmatrix} -2 & 1 \\ -1 & 1 \\ 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} -\frac{9}{5} \\ \frac{4}{5} \end{bmatrix}\right|\right|$$

$$= \left|\left|\begin{bmatrix} 4 \\ 3 \\ 1 \\ -1 \\ -3 \end{bmatrix} - \begin{bmatrix} 4.4 \\ 2.6 \\ 0.8 \\ -1 \\ -2.8 \end{bmatrix}\right|\right|$$

$$= \left|\left|\begin{bmatrix} -0.4 \\ 0.4 \\ 0.2 \\ 0 \\ -0.2 \end{bmatrix}\right|\right|$$

$$= \frac{\sqrt{10}}{5}$$

$$\approx 0.6325$$

(ii)

Suppose the quadratic function $g(x) = n_1x^2 + n_2x + n_3$.

Suppose $B = \begin{bmatrix} 4 & -2 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 4 \\ 3 \\ 1 \\ -1 \\ -3 \end{bmatrix}$, $\mathbf{n} = \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix}$.

Then the least-squares solution of $B\mathbf{n} = \mathbf{b}$ gives the best fits.

$$
\begin{aligned}
\hat{\mathbf{n}} &= (B^TB)^{-1}B^T\mathbf{b}\\
&= \left(\begin{bmatrix} 4 & 1 & 0 & 1 & 4\\ -2 & -1 & 0 & 1 & 2\\ 1 & 1 & 1 & 1 & 1\end{bmatrix}\begin{bmatrix} 4 & -2 & 1\\ 1 & -1 & 1\\ 0 & 0 & 1\\ 1 & 1 & 1\\ 4 & 2 & 1\end{bmatrix}\right)^{-1}\begin{bmatrix} 4 & 1 & 0 & 1 & 4\\ -2 & -1 & 0 & 1 & 2\\ 1 & 1 & 1 & 1 & 1\end{bmatrix}\begin{bmatrix} 4\\ 3\\ 1\\ -1\\ -3\end{bmatrix}\\
&= \begin{bmatrix} 34 & 0 & 10\\ 0 & 10 & 0\\ 10 & 0 & 5\end{bmatrix}^{-1}\begin{bmatrix} 4 & 1 & 0 & 1 & 4\\ -2 & -1 & 0 & 1 & 2\\ 1 & 1 & 1 & 1 & 1\end{bmatrix}\begin{bmatrix} 4\\ 3\\ 1\\ -1\\ -3\end{bmatrix}\\
&= \begin{bmatrix} \frac{1}{14} & 0 & -\frac{1}{7}\\ 0 & \frac{1}{10} & 0\\ -\frac{1}{7} & 0 & \frac{17}{35}\end{bmatrix}\begin{bmatrix} 4 & 1 & 0 & 1 & 4\\ -2 & -1 & 0 & 1 & 2\\ 1 & 1 & 1 & 1 & 1\end{bmatrix}\begin{bmatrix} 4\\ 3\\ 1\\ -1\\ -3\end{bmatrix}\\
&= \begin{bmatrix} \frac{1}{7} & -\frac{1}{14} & -\frac{1}{7} & -\frac{1}{14} & \frac{1}{7}\\ -\frac{1}{5} & -\frac{1}{10} & 0 & \frac{1}{10} & \frac{1}{5}\\ -\frac{3}{35} & \frac{12}{35} & \frac{17}{35} & \frac{12}{35} & -\frac{3}{35}\end{bmatrix}\begin{bmatrix} 4\\ 3\\ 1\\ -1\\ -3\end{bmatrix}\\
&= \begin{bmatrix} -\frac{1}{7}\\ -\frac{9}{5}\\ \frac{38}{35}\end{bmatrix}
\end{aligned}
$$

$g(x) = -\frac{1}{7}x^2 - \frac{9}{5}x + \frac{38}{35}$.

$$
\begin{aligned}
E &= ||\mathbf{b} - B\hat{\mathbf{n}}||\\
&= ||\begin{bmatrix} 4\\ 3\\ 1\\ -1\\ -3\end{bmatrix} - \begin{bmatrix} 4 & -2 & 1\\ 1 & -1 & 1\\ 0 & 0 & 1\\ 1 & 1 & 1\\ 4 & 2 & 1\end{bmatrix}\begin{bmatrix} -\frac{1}{7}\\ -\frac{9}{5}\\ \frac{38}{35}\end{bmatrix}||\\
&= ||\begin{bmatrix} 4\\ 3\\ 1\\ -1\\ -3\end{bmatrix} - \begin{bmatrix} \frac{144}{35}\\ \frac{96}{35}\\ \frac{38}{35}\\ -\frac{6}{7}\\ -\frac{108}{35}\end{bmatrix}||\\
&= ||\begin{bmatrix} -\frac{4}{35}\\ \frac{9}{35}\\ -\frac{3}{35}\\ -\frac{1}{7}\\ \frac{3}{35}\end{bmatrix}||\\
&= \frac{2\sqrt{35}}{35}\\
&\approx 0.3381
\end{aligned}
$$

exercise 11

1. When

(a) $Q(\mathbf{x}) = -2x_1^2 - x_2^2 + 4x_1x_2 + 4x_2x_3$,

(b) $Q(\mathbf{x}) = 7x_1^2 + x_2^2 + 7x_3^2 - 8x_1x_2 - 4x_1x_3 - 8x_2x_3$,

find a unit vector $\boldsymbol{x}$ in $\mathbb{R}^3$ at which $Q(\boldsymbol{x})$ is maximized, subject to $\boldsymbol{x}^T\boldsymbol{x} = 1$.

solution:

(a) $A = \begin{bmatrix} -2 & 2 & 0 \\ 2 & -1 & 2 \\ 0 & 2 & 0 \end{bmatrix} = PDP^{-1}$

$$det(A - \lambda I) = 0 \Rightarrow \lambda_{max} = 2 : \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{bmatrix}$$

$$\mathbf{x} = P\mathbf{e_1} = \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{bmatrix}.$$

(b) $A = \begin{bmatrix} 7 & -4 & -2 \\ -4 & 1 & -4 \\ -2 & -4 & 7 \end{bmatrix} = PDP^{-1}$

$$det(A - \lambda I) = 0 \Rightarrow \lambda_{max} = 9 : \begin{bmatrix} \frac{-2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ 0 \end{bmatrix}$$

$$\mathbf{x} = P\mathbf{e_1} = \begin{bmatrix} \frac{-2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ 0 \end{bmatrix}.$$

2. Find the SVD of $A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$.

3. (a) Suppose A is square and invertible. Find a singular value decomposition of A^{-1}.

(b) Show that the columns of V are eigenvectors of A^TA, the columns of U are eigenvectors of AA^T, and the diagonal entries of Σ are the singular values of A.

4. (a) Show that if T and U are positive semidefinite $n \times n$ matrices such that $T^2 = U^2$, then $T = U$.

(b) A is a positive definite matrix and $A = U\Sigma V^T$ is a singular value decomposition of A. Prove that $U = V$.

solution:

(a) To prove $T = U$, we only need to prove $T\mathbf{x} = U\mathbf{x}$ for any $\mathbf{x} \in \mathbb{R}^n$.

T is positive semidefinite matrix $\Rightarrow Q(\mathbf{x}) = \mathbf{x}^T T\mathbf{x}$ exists $\Rightarrow T$ is symmetric $\Rightarrow T$ is orthogonally diagonalizable

$T = PDP^T$, where $P = [\,\mathbf{p_1} \quad \mathbf{p_2} \quad \cdots \quad \mathbf{p_n}\,]$.

$\mathbf{p_1},\ \mathbf{p_2},\ \cdots,\ \mathbf{p_n}$ are n linearly independent eigenvectors that correspond to eigenvalues $\lambda_1,\ \lambda_2,\ \cdots,\ \lambda_n$,

where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$.

for $1 \leq k \leq n$,

$U^2\mathbf{p_k} = T^2\mathbf{p_k} = \lambda_k^2\mathbf{p_k}$.

(i) $\lambda_k = 0$:

$(U\mathbf{p_k})^T(U\mathbf{p_k}) = \mathbf{p_k}^T U^2 \mathbf{p_k} = 0 \Rightarrow U\mathbf{p_k} = 0 = T\mathbf{p_k}$.

(ii) $\lambda_k > 0$:

$U^2\mathbf{p_k} = \lambda_k^2 \mathbf{p_k} \Rightarrow (U^2 - \lambda_k^2 I)\mathbf{p_k} = 0 \Rightarrow (U + \lambda_k I)(U - \lambda_k I)\mathbf{p_k} = 0$.

If $det(U + \lambda_k I) = 0$, then $-\lambda_k$ is an eigenvalue of U, which is impossible because U is positive semidefinite.

$\Rightarrow det(U + \lambda_k I) \neq 0 \Rightarrow U + \lambda_k I$ is invertible.

$(U - \lambda_k I)\mathbf{p_k} = 0 \Rightarrow U\mathbf{p_k} = \lambda_k \mathbf{p_k} = T\mathbf{p_k}$.

$\forall\, \mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} = c_1\mathbf{p_1} + c_2\mathbf{p_2} + \cdots + c_n\mathbf{p_n}$.

$$\begin{aligned} T\mathbf{x} &= T(c_1\mathbf{p_1} + c_2\mathbf{p_2} + \cdots + c_n\mathbf{p_n}) \\ &= c_1 T\mathbf{p_1} + c_2 T\mathbf{p_2} + \cdots + c_n T\mathbf{p_n} \\ &= c_1 U\mathbf{p_1} + c_2 U\mathbf{p_2} + \cdots + c_n U\mathbf{p_n} \\ &= U(c_1\mathbf{p_1} + c_2\mathbf{p_2} + \cdots + c_n\mathbf{p_n}) \\ &= U\mathbf{x} \end{aligned}$$

$\Rightarrow T = U$.

(b)

$A = U\Sigma V^T$.

A is a positive definite matrix $\Rightarrow$ A is symmetric. Then,

$$\begin{aligned} A^2 = A^T A &= V\Sigma U^T U \Sigma V^T \\ &= V\Sigma^2 V^T \\ &= (V\Sigma V^T)^2 \end{aligned}$$

By part (a),

$A = V\Sigma V^T \Rightarrow U = V$.

group 1 work 1 4

Group 1 work

1. When

(a) $Q(\mathbf{x}) = -2x_1^2 - x_2^2 + 4x_1x_2 + 4x_2x_3$,

(b) $Q(\mathbf{x}) = 7x_1^2 + x_2^2 + 7x_3^2 - 8x_1x_2 - 4x_1x_3 - 8x_2x_3$,

find a unit vector $\boldsymbol{x}$ in $\mathbb{R}^3$ at which $Q(\boldsymbol{x})$ is maximized, subject to $\boldsymbol{x}^T\boldsymbol{x} = 1$.

solution:

(a) $A = \begin{bmatrix} -2 & 2 & 0 \\ 2 & -1 & 2 \\ 0 & 2 & 0 \end{bmatrix} = PDP^{-1}$

$$det(A - \lambda I) = 0 \Rightarrow \lambda_{max} = 2 : \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{bmatrix}$$

$$\mathbf{x} = P\mathbf{e_1} = \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{bmatrix}.$$

(b) $A = \begin{bmatrix} 7 & -4 & -2 \\ -4 & 1 & -4 \\ -2 & -4 & 7 \end{bmatrix} = PDP^{-1}$

$det(A - \lambda I) = 0 \Rightarrow \lambda_{max} = 9 : \begin{bmatrix} \frac{-2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ 0 \end{bmatrix}$

$\mathbf{x} = P\mathbf{e_1} = \begin{bmatrix} \frac{-2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ 0 \end{bmatrix}$.

4. (a) Show that if T and U are positive semidefinite $n \times n$ matrices such that $T^2 = U^2$, then $T = U$.

(b) A is a positive definite matrix and $A = U\Sigma V^T$ is a singular value decomposition of A. Prove that $U = V$.

solution:

(a) To prove $T = U$, we only need to prove $T\mathbf{x} = U\mathbf{x}$ for any $\mathbf{x} \in \mathbb{R}^n$.

T is positive semidefinite matrix $\Rightarrow Q(\mathbf{x}) = \mathbf{x}^T T\mathbf{x}$ exists $\Rightarrow T$ is symmetric $\Rightarrow T$ is orthogonally diagonalizable

$T = PDP^T$, where $P = [\,\mathbf{p_1} \quad \mathbf{p_2} \quad \cdots \quad \mathbf{p_n}\,]$.

$\mathbf{p_1},\ \mathbf{p_2},\ \cdots,\ \mathbf{p_n}$ are n linearly independent eigenvectors that correspond to eigenvalues $\lambda_1,\ \lambda_2,\ \cdots,\ \lambda_n$,

where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$.

for $1 \leq k \leq n$,

$U^2\mathbf{p_k} = T^2\mathbf{p_k} = \lambda_k^2\mathbf{p_k}$.

(i) $\lambda_k = 0$:

$(U\mathbf{p_k})^T(U\mathbf{p_k}) = \mathbf{p_k}^T U^2\mathbf{p_k} = 0 \Rightarrow U\mathbf{p_k} = 0 = T\mathbf{p_k}$.

(ii) $\lambda_k > 0$:

$U^2\mathbf{p_k} = \lambda_k^2\mathbf{p_k} \Rightarrow (U^2 - \lambda_k^2 I)\mathbf{p_k} = 0 \Rightarrow (U + \lambda_k I)(U - \lambda_k I)\mathbf{p_k} = 0$.

If $det(U + \lambda_k I) = 0$, then $-\lambda_k$ is an eigenvalue of U, which is impossible because U is positive semidefinite.

$\Rightarrow det(U + \lambda_k I) \neq 0 \Rightarrow U + \lambda_k I$ is invertible.

$(U - \lambda_k I)\mathbf{p_k} = 0 \Rightarrow U\mathbf{p_k} = \lambda_k\mathbf{p_k} = T\mathbf{p_k}$.

$\forall\, \mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} = c_1\mathbf{p_1} + c_2\mathbf{p_2} + \cdots + c_n\mathbf{p_n}$.

$$\begin{aligned} T\mathbf{x} &= T(c_1\mathbf{p_1} + c_2\mathbf{p_2} + \cdots + c_n\mathbf{p_n}) \\ &= c_1T\mathbf{p_1} + c_2T\mathbf{p_2} + \cdots + c_nT\mathbf{p_n} \\ &= c_1U\mathbf{p_1} + c_2U\mathbf{p_2} + \cdots + c_nU\mathbf{p_n} \\ &= U(c_1\mathbf{p_1} + c_2\mathbf{p_2} + \cdots + c_n\mathbf{p_n}) \\ &= U\mathbf{x} \end{aligned}$$

$\Rightarrow T = U$.

(b)

$A = U\Sigma V^T$.

A is a positive definite matrix $\Rightarrow$ A is symmetric. Then,

$$
\begin{aligned}
A^2 = A^T A &= V\Sigma U^T U \Sigma V^T \\
&= V\Sigma^2 V^T \\
&= (V\Sigma V^T)^2
\end{aligned}
$$

By part (a),

$A = V\Sigma V^T \Rightarrow U = V$.

草稿

$$A = \begin{bmatrix} 1-\lambda & -2 & 4 \\ 2 & 3-\lambda & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

when $\lambda = 0,\ A = \begin{bmatrix} 1 & -2 & 4 \\ 2 & 3 & 1 \\ 1 & 1 & 1 \end{bmatrix}, A - I = \begin{bmatrix} 0 & -2 & 4 \\ 2 & 2 & 1 \\ 1 & 1 & 0 \end{bmatrix}$

$AB + I = A^2 + B$

$(A-I)B = (A-I)(A+I)$

$det(A-I) = -2 \neq 0 \Rightarrow A - I$ is invertible

$B = A + I$