SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

Matrix decompositions in linear algebra

av

Joakim Berg

2014 - No 9

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET, 106 91 STOCKHOLM

Matrix decompositions in linear algebra

Joakim Berg

Självständigt arbete i matematik 15 högskolepoäng, Grundnivå

Handledare: Yishao Zhou

2014

Matrix decompositions in linear algebra

Joakim Berg

April 29, 2014

Thanks to
Jörgen Backelin for encouraging me when I only had a hunch,
Yishao Zhou for helping me explore the world of matrices and to
Alex Loiko for trying to understand my sometimes not so structured thoughts.

Abstract

This paper is about exploring matrix decompositions in different mathematical topics. By mainly using Gauss-elimination we can solve problems such as determining an orthogonal basis, Jordan chains and the Jordan decomposition, the construction of a feedback matrix to reach the desired eigenvalues. This paper is intended to provide a new way of thinking in solving many different mathematical problems.

Contents

Chapter 1

Introduction

The idea of this paper came when I sat in the classroom to listening a lecture on how to do the Gram-Schmidt process in $\mathbb{R}^n$ and I though to myself, there must be a better way to do this. And there was! I found out that you could use Gauss elimination to do the same thing(this method is explained in 2.3). And then I started to think. What else can you do only using Gauss elimination?

I started to explore different kinds of matrix decompositions and linear algebra problems with this approach. I limited myself to only use methods that involved variations of Gauss elimination and matrix multiplication. I found that a lot of the problems in linear algebra could explain in terms of matrix decompositions. In this piper I'm going to show how to look at linear algebra almost entirely in terms of matrix decompositions.

1.1 Matrices in linear algebra

Matrices are an important part of linear algebra. In this section we shall introduce different notations used in matrix theory. Many linear relations can be written in a compact way using matrices. I shall give some examples to show how matrices naturally appear in many objects after introducing some basic and conventional mathematical notations. I assume that the reader is familiar with the basic concepts on linear spaces, also called vector spaces, a basis in a vector space, linear (in)dependency of vectors, dimension of a subspace, linear transformation and so on, (see for example [1,2]).

Let $\mathbb{K}$ be a field and $\mathbb{K}^{n\times m}$ be the set of all $n \times m$ (n rows, m columns) matrices where every element of the matrix is in $\mathbb{K}$. Denote by $\mathbb{K}^n = \mathbb{K}^{n\times 1}$ the set of all (column) vectors with n dimensions. As usual I will denote by $\mathbb{R}$ the real numbers and by $\mathbb{C}$ the complex ones.

A very simple example for writing an object in matrix form is a linear combination of a set of vectors $b_1, b_2, ..., b_k \in \mathbb{R}^n$: $\lambda_1 b_1 + \lambda_2 b_2 + \cdots \lambda_k b_k$ where $\lambda_1, ...\lambda_k \in \mathbb{R}$. In matrix form we have

$$\begin{pmatrix} b_1 & b_2 & \cdots & b_k \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_k \end{pmatrix} = \lambda_1 b_1 + \lambda_2 b_2 + \cdots \lambda_k b_k.$$

A second familiar example is a system of linear equations

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m = b_n \end{cases}$$

This can be written in the matrix form

$$AX = B$$

where

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & 2_{2m} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

A third example is the connection between polynomials and matrices. This connection is both by the characteristic polynomial and, as we shall see later in the paper, by vectors. The matrix under demonstrate both connection to polynomials.

$$C_q^\sharp = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 & -q_0 \\ 1 & 0 & \cdots & 0 & 0 & -q_1 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & -q_{n-2} \\ 0 & 0 & \cdots & 0 & 1 & -q_{n-1} \end{pmatrix}$$

First we can see that this matrix has the characteristic polynomial of this matrix is

$$q(z) = z^n + q_{n-1}z^{n-1} + ... + q_0$$

we can prove this by assuming

$$\begin{vmatrix} z & 0 & \cdots & 0 & 0 & q_1 \\ -1 & z & \cdots & 0 & 0 & q_2 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & z & q_{n-2} \\ 0 & 0 & \cdots & 0 & -1 & z+q_{n-1} \end{vmatrix} = z^{n-1} + q_{n-1}z^{n-2} + \cdots + q_2 z + q_1$$

Now expand along the first row we obtain

$$\begin{aligned} &\begin{vmatrix} z & 0 & \cdots & 0 & 0 & q_0 \\ -1 & z & \cdots & 0 & 0 & q_1 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & z & q_{n-2} \\ 0 & 0 & \cdots & 0 & -1 & z+q_{n-1} \end{vmatrix} \\ =&z\begin{vmatrix} z & 0 & \cdots & 0 & 0 & q_1 \\ -1 & z & \cdots & 0 & 0 & q_2 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & z & q_{n-2} \\ 0 & 0 & \cdots & 0 & -1 & z+q_{n-1} \end{vmatrix} + (-1)^{1+n}q_0 \underbrace{\begin{vmatrix} -1 & z & 0 & \cdots & 0 & 0 \\ 0 & -1 & z & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & z \\ 0 & 0 & 0 & \cdots & 0 & -1 \end{vmatrix}}_{n-1} \\ =&z(z^{n-1} + q_{n-1}z^{n-2} + \cdots + q_2 z + q_1) + (-1)^{n+1}(-1)^{n-1}q_0 \\ =&z^n + q_{n-1}z^{n-1} + \cdots + q_2 z^2 + q_1 z + q_0 \end{aligned}$$

The connection with vectors has to do with to polynomial division. Consider the polynomial $a(z) = a_{n-1}z^{n-1} + ... + a_0$. Now if we take $za(z)$ and do polynomial division with $q(z)$ we get that the reminder of non negative power is the same as if we take $C_q^\sharp a$ where $a = \begin{pmatrix} a_0 \\ \vdots \\ a_{n-1} \end{pmatrix}$.

The central topic of this paper is on different kinds of matrix decompositions used in some mathematical disciplines such as study of structure of linear transformations, numerical linear algebra, mathematical control theory, to mention a few. The main idea is to perform Gauss elimination in decompositions of matrices. The purpose is to look at many existing topics from a new angle. It turns out that the treatment on topics in finding feedback matrix in this paper lead a result seemed to be new, at least in its explicit form and characterization.

1.2 Definitions

In this section I collect notations and definitions used frequently in the sequel. Most conventions are from the references given in the end of the paper.

Definition 1 *The transpose of a matrix* $A \in \mathbb{K}^{n\times m}$ *is denoted* A^T *and has A:s columns as rows.*

Definition 2 *The identity looks like* $\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$ *an is denoted as* I_n *if it is an* $n \times n$ *matrix. If nothing else is said* I *is the Identity matrix with the right size.*

Definition 3 *The inverse of a matrix* $A \in \mathbb{K}^{n\times n}$ *is denoted as* A^{-1} *and has the property that* $AA^{-1} = A^{-1}A = I_n$

Definition 4 *The image of a matrix* $A \in \mathbb{K}^{n\times m}$ *is denoted*
$\mathrm{Im}(A) = \{Ax|x \in \mathbb{K}^m\}$

Definition 5 *The kernel of a matrix* $A \in \mathbb{K}^{n\times m}$ *is denoted*
$\mathrm{Ker}(A) = \{x|Ax = 0\}$

Definition 6 *A full rank* $A \in \mathbb{K}^{n\times m}$ *is a matrix where*
$\mathrm{Ker}(A) = 0$ *or* $\mathrm{Ker}(A^T) = 0$.

(Note: there are other definitions of full rank but this one is the one I find most suitable for this paper.)

Definition 7 *For a full rank matrix* $K \in \mathbb{R}^{n\times m}$ *and* $n \geq m$ *the matrix* $K^\dagger$ *will be denoted as* $K^\dagger = (K^TK)^{-1}K^T$ *and if* $n \leq m$ *then* $K^\dagger = K^T(KK^T)^{-1}$

Note that I shall write 0 for the zero matrix of appropriate size according to the context, that is I do not, in general, specify the dimension of the zero matrix for simplicity.

1.3 Block matrices

I shall use block matrices very often. Usually we obtain them from ordinary matrices by dividing then by several horizontal and/or vertical lines into block. For example

$$C_q^\sharp = \left(\begin{array}{ccccc|c} 0 & 0 & \cdots & 0 & 0 & -q_0 \\ \hline 1 & 0 & \cdots & 0 & 0 & -q_1 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & -q_{n-2} \\ 0 & 0 & \cdots & 0 & 1 & -q_{n-1} \end{array}\right)$$

We divide $C_q^\sharp$ into four blocks

$$C_q^\sharp = \begin{pmatrix} X & Y \\ U & W \end{pmatrix}$$

with

$$X = \underbrace{\begin{pmatrix} 0 & 0 & \cdots & 0 \end{pmatrix}}_{n-1}, \quad Y = -q_0, \quad U = I_{n-1}, \quad W = \begin{pmatrix} -q_0 \\ -q_1 \\ \vdots \\ -q_{n-1} \end{pmatrix}$$

or likewise

$$C_q^\sharp = \left(\begin{array}{ccccc|c} 0 & 0 & \cdots & 0 & 0 & -q_0 \\ 1 & 0 & \cdots & 0 & 0 & -q_1 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & -q_{n-2} \\ \hline 0 & 0 & \cdots & 0 & 1 & -q_{n-1} \end{array}\right) = \begin{pmatrix} X' & Y' \\ U' & W' \end{pmatrix}$$

with

$$X' = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}, Y' = \begin{pmatrix} -q_0 \\ \vdots \\ -q_{n-2} \end{pmatrix}, U' = \underbrace{\begin{pmatrix} 0 & \cdots & 0 & 1 \end{pmatrix}}_{n-1}, W' = -q_{n-1}$$

When working on multiplication matrices we have to divide the matrix blocks into right sizes so that multiplication makes sense. The transpose of a block works similar to transpose of an ordinary matrix but it is important to transpose each block, e.g.

$$(C_q^\sharp)^T = \begin{pmatrix} X^T & U^T \\ Y^T & W^T \end{pmatrix} = \begin{pmatrix} X'^T & U'^T \\ Y'^T & W'^T \end{pmatrix}.$$

Proposition 1 *Assume that A and B are square matrices. Then*

$$\begin{vmatrix} A & 0 \\ C & B \end{vmatrix} = \det(A)\det(B).$$

Proof. If A or B is singular the equality is clearly true, for the right hand side will be zero (either $\det(A) = 0$ or $\det(B) = 0$). But the left hand side will also be zero becasue either the first row block consists of linearly dependent row or the first column block consists of linearly dependent columns, which lead to a zero determinant.

Now we assume that either A or B is nonsigular. Observe that

$$\begin{pmatrix} A & 0 \\ C & B \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} I & 0 \\ B^{-1}C & I \end{pmatrix}$$

Hence $\begin{vmatrix} A & 0 \\ C & B \end{vmatrix} = \begin{vmatrix} A & 0 \\ 0 & I \end{vmatrix} \begin{vmatrix} I & 0 \\ 0 & B \end{vmatrix} \begin{vmatrix} I & 0 \\ B^{-1}C & I \end{vmatrix} =$
$\det(A)\det(I)\det(I)\det(B)\det(I)\det(I) = \det(A)\det(B)$

Proposition 2 *Assume that A is a nonsigular matrix. Then*

$$\begin{vmatrix} A & D \\ C & B \end{vmatrix} = \det(A)\det(B - CA^{-1}D).$$

Similarly if B is nonsigular,

$$\begin{vmatrix} A & D \\ C & B \end{vmatrix} = \det(B)\det(A - DB^{-1}C).$$

where A, B, C, D are of appropriate dimension.

Proof. Observe that (by Gause elimination blockwise) assuming A is nonsingular,

$$\begin{pmatrix} I & 0 \\ -CA^{-1} & I \end{pmatrix} \begin{pmatrix} A & D \\ C & B \end{pmatrix} = \begin{pmatrix} A & D \\ 0 & B - CA^{-1}D \end{pmatrix}$$

Then

$$\begin{vmatrix} I & 0 \\ -CA^{-1} & I \end{vmatrix} \begin{vmatrix} A & D \\ C & B \end{vmatrix} = \begin{vmatrix} A & D \\ 0 & B - CA^{-1}D \end{vmatrix}$$

which is by the property that the determinant of a matrix is equal to the determinants of its transpose and Proposition 1

$$\begin{vmatrix} A & D \\ C & B \end{vmatrix} = \begin{vmatrix} A & D \\ 0 & B - CA^{-1}D \end{vmatrix} = \det(A)\det(B - CA^{-1}D)$$

as desired.

Note that the property $\det(AB) = \det(A)\det(B)$ used in the proofs requires that A and B be square matrices but this does not hold if they are non-square. However we have the flowing important theorem.

Proposition 3 *Let A be $n \times m$ and B be $m \times n$, then*

$$\det(I_n - AB) = \det(I_m - BA).$$

In particular, if $m = 1$ then

$$\det(I_n - AB) = 1 - BA$$

Proof. Compute the determinant $\begin{vmatrix} I_n & A \\ B & I_m \end{vmatrix}$ using the previous proposition.

$$\begin{vmatrix} I_n & A \\ B & I_m \end{vmatrix} = \det(I_n)\det(I_m - BI_n^{-1}A) = \det(I_m - BA)$$

On the other hand,

$$\begin{vmatrix} I_n & A \\ B & I_m \end{vmatrix} = \det(I_m)\det(I_n - AI_m^{-1}B) = \det(I_n - AB).$$

Thus $\det(I_n - AB) = \det(I_m - BA)$.

Clearly if $m = 1$, A is a columne vector and B is a row vector. Hence $I_m - BA$ is a scalar and equals $1 - BA$. Therefore, $\det(I_n - AB) = 1 - BA$.

Chapter 2

Matrix decompositions

In this chapter I will explain how to do different decompositions. I will do these decompositions by using Gauss and Gauss-Jordan elimination and different variants of those.

2.1 Basic Theory

As I mentioned the first thing you have to know is how to use Gauss elimination to compute the inverse of a given matrix. Let $A \in \mathbb{K}^{n\times n}$ be a nonsingular matrix. As we do in our linear algebra class, I augment the matrix A with the identity matrix $I = I_n$ as $(A \mid I)$. Then we do row operations on this augmented matrix until the matrix in the position of A becomes I. Call the matrix on the right C. Then C is the inverse of A, i.e. $AC = CA = I$. This procedure is called Gauss-Jordan elimination. For example , $A = \begin{pmatrix} 1 & 1 & -2 \\ 2 & 0 & 2 \\ -1 & 0 & 2 \end{pmatrix}$. Now we perform Gauss-Jordan elimination on

$$\left(\begin{array}{ccc|ccc} 1 & 1 & -2 & 1 & 0 & 0 \\ 2 & 0 & 2 & 0 & 1 & 0 \\ -1 & 0 & 2 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 1 & -2 & 1 & 0 & 0 \\ 0 & -2 & 6 & -2 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{array}\right) \sim$$

$$\left(\begin{array}{ccc|ccc} 1 & 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & -2 & 6 & -2 & 1 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 6 & 0 & 1 & 2 \end{array}\right) \sim$$

$$\left(\begin{array}{ccc|ccc} 1 & 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & \frac{1}{6} & \frac{1}{3} \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & \frac{1}{3} & -\frac{1}{3} \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & \frac{1}{6} & \frac{1}{3} \end{array}\right)$$

Now we have

$$A^{-1} = \begin{pmatrix} 0 & \frac{1}{3} & -\frac{1}{3} \\ 1 & 0 & 1 \\ 0 & \frac{1}{6} & \frac{1}{3} \end{pmatrix}$$

Note that the process of row reducing until the matrix is reduced, as done above, is sometimes referred to as Gauss-Jordan elimination, to distinguish it from stopping after reaching echelon form. In the above example it is the next last step. By row echelon form of a matrix we mean that the matrix satisfies the following condition ([3]):

- All nonzero rows (rows with at least one nonzero element) are above any rows of all zeroes (all zero rows, if any, belong at the bottom of the matrix).
- The leading coefficient (the first nonzero number from the left, also called the pivot) of a nonzero row is always strictly to the right of the leading coefficient of the row above it.
- All entries in a column below a leading entry are zeroes (implied by the first two criteria).

The aim of doing this example is to make the following point. At each step we have the form

$$(A \mid I) \sim (B \mid C)$$

This is equivalent to

$$CA = B.$$

In fact, performing Gauss elimination on A to get B is to multiply A by C from left, and C consists of the row operations up to this step. Note that this is correct for $A \in \mathbb{K}^{n\times m}$ as well. We shall use them interchangeably in the sequel.

2.1.1 Determination of a basis for a kernel

Now we know how to perform Gauss elimination to find the inverse of the matrix A and the solution is the matrix C when $(A \mid I) \sim (I \mid C)$. Note that we just read off what we have obtained from the last elimination. I claim that this can be used to find a basis of the kernel of a matrix A.

Given a mtrix $A \in \mathbb{K}^{m\times n}$ we can do the following:
Perform Gauss elimination on $(A^T \mid I_n)$ until we have the form

$$\left(\begin{array}{c|c} X & \multirow{2}{*}{C} \\ 0 & \end{array}\right) = \left(\begin{array}{c|c} X & C'' \\ 0 & C' \end{array}\right)$$

i.e. $CA^T = \begin{pmatrix} X \\ 0 \end{pmatrix}$. (Note that $A(C''^T \; C'^T) = (X \; 0)$.) This implies that $AC'^T = 0$. C' gives a basis of $\mathrm{Ker}(A)$: the columns of C'^T. Moreover since X has full rank, we have

$$\mathrm{Ker}(A) = \mathrm{Im}\left(C'^T\right).$$

Example 1 *Take the matrix* $A = \begin{pmatrix} 1 & 2 & 3 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix}$. *Set* $A' = \left(\begin{array}{cc|cccc} 1 & 1 & 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 & 0 \\ 3 & 1 & 0 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 0 & 1 \end{array}\right)$.

Now we can do Gauss-elimination:

$$\left(\begin{array}{cc|cccc} 1 & 1 & 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 & 0 \\ 3 & 1 & 0 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{cc|cccc} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & -1 & -2 & 1 & 0 & 0 \\ 0 & -2 & -3 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{cc|cccc} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & -1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & -3 & 1 & 0 & 1 \end{array}\right).$$

We take out the last rows: $\begin{pmatrix} 1 & -2 & 1 & 0 \\ -3 & 1 & 0 & 1 \end{pmatrix}$.

$\begin{pmatrix} 1 & 2 & 3 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & -3 \\ -2 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = 0$, *as expected. This gives us that*

$\mathrm{Ker}(A) = \{ \begin{pmatrix} 1 & -3 \\ -2 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} x | x \in \mathbb{R}^2 \}$.

2.1.2 Determination of the intersection of images of two matrices

Another thing we can do is to find a basis for $\mathrm{Im}(N) \cap \mathrm{Im}(K)$ where N, K are $n \times m$ matrices. This is not as trivial as to find a basis in the kernel of a matrix. However as we shall see it turns out to the same problem we have to deal with. There are other methods to do this, but I'm going to use one where we also can find a vector space of as big rank as possible in $\mathrm{Im}(N) \setminus \mathrm{Im}(K) \setminus \{0\}$.

We want to find all linearly independent solutions x and y such that $Nx = Ky$. That is, x, y is a solution of $(N \; -K) \begin{pmatrix} x \\ y \end{pmatrix} = 0$. Now we can apply the method for finding the kernel to this problem. Do Gauss elimination on this matrix augmented with I_{2m} until we get the form we need, i.e.

$$\left(\begin{array}{c|c} N^T & \\ \hline -K^T & I_{2m} \end{array}\right) \sim \left(\begin{array}{c|cc} D & A & 0 \\ \hline D' & B_1 & C_1 \\ 0 & B_2 & C_2 \end{array}\right)$$

That is,

$$\left(\begin{array}{cc} A & 0 \\ \hline B_1 & C_1 \\ B_2 & C_2 \end{array}\right) \begin{pmatrix} N^T \\ -K^T \end{pmatrix} = \begin{pmatrix} D \\ \hline D' \\ 0 \end{pmatrix} \Leftrightarrow \left(\begin{array}{c} AN^T \\ \hline B_1N^T - C_1K^T \\ B_2N^T - C_2K^T \end{array}\right) = \begin{pmatrix} D \\ \hline D' \\ 0 \end{pmatrix}$$

The second block matrix equation is

$$\begin{pmatrix} B_1N^T - C_1K^T \\ B_2N^T - C_2K^T \end{pmatrix} = \begin{pmatrix} D' \\ 0 \end{pmatrix}$$

From this we see that $B_2N^T = C_2K^T$, or equivalently $NB_2^T = KC_2^T$. Hence

$$\mathrm{Im}(NB_2^T) = \mathrm{Im}(KC_2^T) = \mathrm{Im}(N) \cap \mathrm{Im}(K)$$

Then, we have found a basis in $\mathrm{Im}(N) \cap \mathrm{Im}(K)$: the columns of B_2^T or the columns of C_2^T. If there is no zero row below D then the intersection is $\{0\}$.

The above computation clearly shows that

$$\mathrm{Im}(NB_1^T) \cap \mathrm{Im}(K) = \{0\}, \quad \mathrm{Im}(KC_1^T) \cap \mathrm{Im}(N) = \{0\}$$

since $B_1N^T - C_1K^T = D'$, that is $NB_1^T = KC_1^T + D'^T$, or $KC_1^T = NB_1^T - D'^T$ where $D' \neq 0$ by construction. Hence,

$$\mathrm{Im}(NB_1^T) \subset (\mathrm{Im}(N) \setminus \mathrm{Im}(K)) \cup \{0\}, \quad \mathrm{Im}(KC_1^T) \subset (\mathrm{Im}(K) \setminus \mathrm{Im}(N)) \cup \{0\}$$

We can also see that $\mathrm{Im}((N, K)) = \mathrm{Im}((NB_1^T, NC_1^T, NB_2^T, NC_2^T)) =$
$\mathrm{Im}((NB_1^T, NC_1^T, NB_2^T)) =$
$(\mathrm{Im}(N) \setminus \mathrm{Im}(K)) \cup (\mathrm{Im}(K) \setminus \mathrm{Im}(N)) \cup (\mathrm{Im}(N) \cap \mathrm{Im}(K))$ and we can draw the conclusion that $\mathrm{Im}(NB_1^T)$ is a vector-space in $\mathrm{Im}(N) \setminus \mathrm{Im}(K) \setminus \{0\}$ with the biggest possible rank, notice that this rank is $rank(N) - rank(NB_2^T)$.

Example 2 *Consider* $N = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 0 & 1 \end{pmatrix}$ *and* $K = \begin{pmatrix} 1 & 3 \\ 1 & 2 \\ 0 & 3 \\ 1 & 2 \end{pmatrix}$. *We can do Gauss-elimination*

$$\left(\begin{array}{cccc|cccc} 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 3 & 2 & 3 & 2 & 0 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{cccc|cccc} 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & -1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 2 & -3 & 0 & 0 & 1 \end{array}\right)$$

$$\left(\begin{array}{cccc|cccc} 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & 0 & -1 & -1 & 1 & 0 \\ 0 & 0 & -2 & 0 & -3 & -2 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{cccc|cc|cc} 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & -2 & 0 & -1 & -1 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & -2 & -1 & -1 & 1 \end{array}\right)$$

Now we can see that $\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -2 \\ -1 \end{pmatrix} + \begin{pmatrix} 1 & 3 \\ 1 & 2 \\ 0 & 3 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ -1 \\ -3 \\ -1 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \\ 3 \\ 1 \end{pmatrix} = 0$

as we expected. We see obviously that a basis for $\mathrm{Im}(N) \cap \mathrm{Im}(K)$ *is* $\begin{pmatrix} 2 \\ 1 \\ 3 \\ 1 \end{pmatrix}$.

And we can also see that $\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ -2 \\ -1 \end{pmatrix} \in \mathrm{Im}(N) \setminus \mathrm{Im}(K) \setminus \{0\}$

we can however not find a proper basis for this space since $\mathrm{Im}(K) \cup \mathrm{Im}(N) \subseteq \mathrm{Im}(N) \setminus \mathrm{Im}(K) \setminus \{0\}$ *and* $\mathrm{Im}(N, K') \subseteq \mathrm{Im}(N, K)$ *where* $\mathrm{Im}(K') = \mathrm{Im}(K) \cup \mathrm{Im}(N)$. *But to find an basis as big as possible can be archived with this method, and it is important in 4.2.*

We can also prove that:

Theorem 1 *Set two full rank matrices* $K \in \mathbb{K}^{n \times m}$ *and* $M \in \mathbb{K}^{m \times n}$ *where* $n > m$. *Then* $rank(MK) = m - rank(\mathrm{Ker}(M) \cap \mathrm{Im}(K))$

Proof
We can find a nonsingular matrix $H \in \mathbb{K}^{m \times m}$ such that $KH = (N, K')$ where $\mathrm{Im}(N) = \mathrm{Ker}(M) \cap \mathrm{Im}(K)$ and since K has full rank we have, $\mathrm{Im}(K') \cap \mathrm{Im}(N) = \{0\}$ and MK' has full rank. We now get $rank(MK) = rank(MKH) = rank((MN, MK')) = rank((0, MK')) = m - rank(\mathrm{Ker}(M) \cap \mathrm{Im}(K))$

2.2 LU decomposition

The LU factorization [1] is to decompose a matrix into an upper triangle matrix (U) and a lower triangular matrix (L). We can do this by Gauss eliminations on an $n \times n$ matrix A to an upper triangular and then take the inverse of the corresponding Matrix.

[1] more abut The LU Factorization exist in: Matrix Computations third edition, Gene H. Golub,Charles F. Van Loan, The Johns Hopkins University press 1996 3.2

Example 3 *We have the matrix* $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 6 \\ 3 & 3 & 5 \end{pmatrix}$. *Then we can do Gauss-elimination so that we get a triangular form.*

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 3 & 6 & 0 & 1 & 0 \\ 3 & 3 & 5 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -1 & 0 & -2 & 1 & 0 \\ 0 & -3 & -4 & -3 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -1 & 0 & -2 & 1 & 0 \\ 0 & 0 & -4 & 2 & -3 & 1 \end{array}\right)$$

Now we take inverse of $\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 2 & -3 & 1 \end{pmatrix}$ *which is* $\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & 3 & 1 \end{pmatrix}$ *and then we get*

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 6 \\ 3 & 3 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & 0 \\ 0 & 0 & -4 \end{pmatrix}$$

I should point out, if there is a permutation in the row operations we can not always make a perfect triangle.

2.3 QR decomposition

This factorization[2] contains a matrix $Q \in \mathbb{R}^{n \times m}, n \geq m, rank(Q) = m, Q^T Q = I_m$ and a matrix $R \in \mathbb{R}^{m \times m}, rank(R) = m$ which is an upper triangular matrix. Set $D \in \mathbb{R}^{n \times m}, rank(D) = m$. Now you can do the LU on the matrix $A = D^T D$ so that $A = LU$, then you take the diagonal in U and take the diagonal as $\frac{1}{\sqrt{diag}}$ with the rows of L^{-1} and it becomes R^{-1}. Then we have that $Q = DR^{-1}, D = QR$. An example of this is.

Example 4 *Let* $D = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$. *Then*

$D^T D = A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{pmatrix}$. *Then the we do Gauss-elimination:*

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 3 & 0 & 1 & 0 \\ 1 & 3 & 6 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & -1 & 1 & 0 \\ 0 & 2 & 5 & -1 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & -1 & 1 & 0 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right)$$

[2]Other methods to do this factorization can be found in: Matrix Computations third edition, Gene H. Golub,Charles F. Van Loan, The Johns Hopkins University press 1996 5.2

Here $Q = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ *and* $R = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$

Next we show why this works. Since $D \in \mathbb{R}^{n \times m}$ with $n \geq m$ the full rank matrix $A \in \mathbb{R}^{n \times m}, n \geq m$ then $A^T A$ has full rank.
Then set

$$A = \begin{pmatrix} a_{11} & . & . & . & a_{1m} \\ . & & & & . \\ . & & & & . \\ . & & & & . \\ a_{m1} & . & . & . & a_{mm} \end{pmatrix}$$

$$B = \begin{pmatrix} b_{11} & . & . & . & b_{1m} \\ 0 & . & & & . \\ . & . & . & & . \\ . & & . & . & . \\ 0 & . & . & 0 & b_{mm} \end{pmatrix}, b_{ii} > 0$$

$$C = \begin{pmatrix} c_{11} & 0 & . & . & 0 \\ . & . & . & & . \\ . & & . & . & . \\ . & & & . & 0 \\ c_{m1} & . & . & . & c_{mm} \end{pmatrix} c_{ii} = 1$$

where $CA = B$, now set the matrix, $P = \begin{pmatrix} \frac{1}{\sqrt{b_{11}}} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sqrt{b_{22}}} & & \\ \vdots & & \ddots & \\ 0 & & & \frac{1}{\sqrt{b_{mm}}} \end{pmatrix}$ Now we want to show that $PCAC^T P = I_m$. I'm going to show this by considering.

$$\frac{1}{\sqrt{b_{ii}}}\begin{pmatrix}c_{i1} & . & . & . & c_{ii} & 0 & . & . & 0\end{pmatrix}\begin{pmatrix}a_{11} & . & . & . & a_{1m}\\ . & & & & .\\ . & & & & .\\ . & & & & .\\ a_{m1} & . & . & . & a_{mm}\end{pmatrix}\frac{1}{\sqrt{b_{ii}}}\begin{pmatrix}c_{i1}\\ .\\ .\\ .\\ c_{ii}\\ 0\\ .\\ .\\ .\\ 0\end{pmatrix} =$$

$$= \frac{1}{b_{ii}}\begin{pmatrix}0 & . & . & . & 0 & b_{ii} & . & . & b_{in}\end{pmatrix}\begin{pmatrix}c_{i1}\\ .\\ .\\ .\\ c_{ii}\\ 0\\ .\\ .\\ .\\ 0\end{pmatrix} = \frac{1}{b_{ii}} \cdot b_{ii} = 1$$

For $i > j$

$$\frac{1}{\sqrt{b_{ii}}}\begin{pmatrix}c_{i1} & . & . & . & c_{ii} & 0 & . & . & 0\end{pmatrix}\begin{pmatrix}a_{11} & . & . & . & a_{1m}\\ . & & & & .\\ . & & & & .\\ . & & & & .\\ a_{m1} & . & . & . & a_{mm}\end{pmatrix}\frac{1}{\sqrt{b_{jj}}}\begin{pmatrix}c_{j1}\\ .\\ .\\ .\\ c_{jj}\\ 0\\ .\\ .\\ .\\ 0\end{pmatrix} =$$

$$= \frac{1}{\sqrt{b_{ii}}} \cdot \frac{1}{\sqrt{b_{jj}}} \begin{pmatrix} 0 & . & . & . & 0 & b_{ii} & . & . & b_{in} \end{pmatrix} \begin{pmatrix} c_{j1} \\ . \\ . \\ . \\ c_{jj} \\ 0 \\ . \\ . \\ . \\ 0 \end{pmatrix} = \frac{1}{\sqrt{b_{ii}}} \cdot \frac{1}{\sqrt{b_{jj}}} \cdot 0 = 0$$

and since A is symmetric we have the same results for $i < j$.
Now if we set $Q = DC^T P$ and set $R^{-1} = C^T P$, we are done.

2.4 Full Rank decomposition

This is a decomposition you can do on any matrix. If we have an $n \times m$ matrix A, the only thing you have to do is a complete elimination of A and then take the same rows form A at the rows that only have a one and zeros after gauss elimination and multiply from the left to the complete Gauss-eliminated one.

Example 5 *Let* $A = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 2 & 1 & 2 & 1 \\ 4 & 5 & 2 & 3 \end{pmatrix}$. *Do the Gauss elimination.*

$$\left(\begin{array}{cccc|ccc} 1 & 2 & 0 & 1 & 1 & 0 & 0 \\ 2 & 1 & 2 & 1 & 0 & 1 & 0 \\ 4 & 5 & 2 & 3 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{cccc|ccc} 1 & 2 & 0 & 1 & 1 & 0 & 0 \\ 0 & -3 & 2 & -1 & -2 & 1 & 0 \\ 0 & -3 & 2 & -1 & -4 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{cccc|ccc} 1 & 2 & 0 & 1 & 1 & 0 & 0 \\ 0 & -3 & 2 & -1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & -2 & -1 & 1 \end{array}\right)$$

Now take the inverse of $\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -2 & -1 & 1 \end{pmatrix}$ *which is* $\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & 1 & 1 \end{pmatrix}$ *and we get that*

$$\begin{pmatrix} 1 & 2 & 0 & 1 \\ 2 & 1 & 2 & 1 \\ 4 & 5 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & -3 & 2 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 4 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & -3 & 2 & -1 \end{pmatrix}$$

There are a couple of things you can do with this factorization. If we assume $A_1 \in \mathbb{K}^{n \times n}$ is a singular matrix then $A_1 = K_1 M_1$ where K_1, M_1 are full rank matrices. Then we set $M_1 K_1 = A_2$ leading to $A_1^2 = K_1 M_1 K_1 M_1 = K_1 A_2 M_1$. If A_2 is singular we can do rank decomposition so that $A_2 = K_2 M_2$. Then set $M_2 K_2 = A_3$. We see that $A_1^3 = K_1 M_1 K_1 M_1 K_1 M_1 = K_1 A_2 A_2 M_1 = K_1 K_2 M_2 K_2 M_2 M_1 = K_1 K_2 A_3 M_2 M_1$ and so on until A_n has full rank. We

can now define $K'_i = K_1 ... K_i$ and $M'_i = M_i ... M_1$.
What can we do with this now? Well if we assume that A_n is the first invertible matrix. Then we can set $E = K'_{n-1} A_n^{1-n} M'_{n-1}$ and we see that $EA^n = K'_{n-1} A_n^{1-n} M'_{n-1} K'_{n-1} A_n M'_{n-1} = K'_{n-1} A_n^{1-n} A_n^n M'_{n-1} = K'_{n-1} A_n M'_{n-1} = A^n$ and we see that any matrix of the form $B = K'_{n-1} H A_n^{1-n} M'_{n-1}$ where H is a full rank matrix, will have the property $EB = EBE = BE = B$. Now we can see that $G = \{K'_{n-1} H A_n^{1-n} M'_{n-1} | \operatorname{Ker}(H) = 0\}$ is a group under matrix multiplication with the Identity element E.
Moreover we can find $\operatorname{Im}(A^n)$ with this method, and we can also prove that the eigenvalues $\neq 0$ of A_1 is the same as those of A_n. But more of that can be found in the Chapter on Jordan decomposition.

Example 6 *Consider the matrix* $A = \begin{pmatrix} 0 & 0 & 1 & 1 \\ -2 & 2 & 2 & 2 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix}$ *Let us try to Gauss-eliminate this matrix*

$$\left(\begin{array}{cccc|cccc} 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ -2 & 2 & 2 & 2 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{cccc|cccc} 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ -2 & 2 & 2 & 2 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right) \sim$$

$$\sim \left(\begin{array}{cccc|cccc} 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ -2 & 2 & 2 & 2 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 \end{array}\right)$$

and since the $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 1 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$. *We see that*

$$A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 & 1 \\ -2 & 2 & 2 & 2 \\ 1 & 0 & 0 & 1 \end{pmatrix} = K_1 M_1$$

$$A_2 = M_1 K_1 = \begin{pmatrix} 0 & 0 & 1 & 1 \\ -2 & 2 & 2 & 2 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 2 \\ 1 & 0 & 1 \end{pmatrix}$$ *Then we do the*

Full rank factorization on A_2.

$$\left(\begin{array}{ccc|ccc}1&0&1&1&0&0\\0&2&2&0&1&0\\1&0&1&0&0&1\end{array}\right)\sim\left(\begin{array}{ccc|ccc}1&0&1&1&0&0\\0&2&2&0&1&0\\0&0&0&-1&0&1\end{array}\right)$$

and we now see that $A_2 = \begin{pmatrix}1&0\\0&1\\1&0\end{pmatrix}\begin{pmatrix}1&0&1\\0&2&2\end{pmatrix} = K_2M_2$ *and then* $A_3 = M_2K_2 = \begin{pmatrix}1&0&1\\0&2&2\end{pmatrix}\begin{pmatrix}1&0\\0&1\\1&0\end{pmatrix} = \begin{pmatrix}2&0\\2&2\end{pmatrix}$ *and have* $E = K_2'A_3^{-2}M_2' =$
$\begin{pmatrix}1&0&0\\0&1&0\\1&0&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&0\\0&1\\1&0\end{pmatrix}\begin{pmatrix}1&0\\-2&1\end{pmatrix}\frac{1}{4}\begin{pmatrix}1&0&1\\0&2&2\end{pmatrix}\begin{pmatrix}0&0&1&1\\-2&2&2&2\\1&0&0&1\end{pmatrix} =$
$\begin{pmatrix}1&0\\0&1\\1&0\\1&0\end{pmatrix}\frac{1}{4}\begin{pmatrix}1&0&1&2\\-4&4&0&0\end{pmatrix}$

Chapter 3

Non eigenvalue problems

In this chapter I am going to look at problems where I don't need the eigenvalues to solve the problems.

3.1 LS problem

The Least Square[1] or LS problem is the problem where you want to find $min_{x\in\mathbb{R}^n}(|Ax-b|)$ for fixed $A \in \mathbb{R}^{m\times n}, m \geq n$ and $b \in \mathbb{R}^m$, where $|b| = \sqrt{b^T b}$. In this section I'm going to show two ways to do this.

3.1.1 QR solution

For an orthogonal $n \times n$ matrix Q we have that $|v| = |Qv|$ for $v \in \mathbb{R}^n$. We can use this to minimize $|Ax - b|$. First we do the QR factorization on A then we take out a basis for the null space of A^T say N and then we do the QR factorization on N^T. So we have that $A = Q_A R_A, N = Q_N R_N$. Set $Q = \begin{pmatrix} Q_A^T \\ Q_N^T \end{pmatrix}$. Now we get that $|Ax - b| = |QAx - Qb| = |\begin{pmatrix} Q_A^T Ax \\ Q_N^T Ax \end{pmatrix} - \begin{pmatrix} Q_A^T b \\ Q_N^T b \end{pmatrix}| = |\begin{pmatrix} Q_A^T Ax - Q_A^T b \\ 0 - Q_N b \end{pmatrix}|$. Let now $x = R_a^{-1} Q_A^T b$. We see that $|Ax - b| = |\begin{pmatrix} Q_A^T B - Q_A^T b \\ -Q_N^T b \end{pmatrix}| = |\begin{pmatrix} 0 \\ Q_N^T b \end{pmatrix}| = |Q_N b|$. This is the best method to actually find out the value of $min_{x\in\mathbb{R}^n}(|Ax - b|) = |Q_N^T b|$.

[1]More about this in: Matrix Computations third edition, Gene H. Golub,Charles F. Van Loan, The Johns Hopkins University press 1996 5.3

3.1.2 The matrix $A^\dagger$

This method is the best method to find out x. The answer to this is $x = (A^TA)^{-1}A^Tb$ we can verify this by checking:
$(R_A^TQ_A^TQ_AR_A)^{-1}A^Tb = (R_A^TR_A)^{-1}A^Tb = R_A^{-1}R_A^{T^{-1}}A^Tb = (A^TA)^{-1}A^Tb$

3.1.3 $||AX - B||$

This is the problem where we shall minimize $||AX - B||$ where $||AX - B||$ is the maximum of $|(AX - B)v|$ where $|v| = 1$. The first thing we can do is to rank factorize $A = KM$ and then set $X = M^\dagger X'$. Now $AX - B = KX' - B$ where K is a tall full rank matrix.
Then we can say that $X' = (x_1, ..., x_m)$ for $x_i \in \mathbb{R}^k$ and $B = (b_1, ..., b_m)$ now we can see that $Kx_i = b_i$ and we can see that x_i is $x_i = K^T(KK^T)^{-1}b_i$ and $X' = (x_1, ..., x_m) = (K^T(KK^T)^{-1}b_1, ..., K^T(KK^T)^{-1}b_m) = K^T(KK^T)^{-1}(b_1, ..., b_m) = K^T(KK^T)^{-1}B$ and we get $X = M^\dagger K^\dagger B$.
This is a solution since for every vector $v \in Im(B)$ will have the solution $x = M^\dagger K^\dagger v$ for minimizing $|Ax - v|$.

3.2 Hessenberg decomposition

The matrix in the following form

$$\begin{pmatrix} * & * & \cdots & * & * \\ * & * & \cdots & * & * \\ 0 & \ddots & & \vdots & \vdots \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & * & * \end{pmatrix}$$

is called a Hessenberg matrix, that is all elements in the matrix below the first off-diagonal line are zero.

Now we use Gauss elimination to reduce any matrix to the Hessenberg form, in the sense of a similarity transform. Note that it is not the same as the Hessenberg decomposition in numerical literature where often it requires the transformation matrix be to orthogonal (unitary). Why I am interested in this decomposition will become apparent later.

This decomposition[2] is to find an matrix U such that $UAU^{-1} = \begin{pmatrix} * & * & \cdots & & * \\ * & * & & & \vdots \\ 0 & \ddots & & & \\ \vdots & \ddots & & & \\ 0 & \cdots & 0 & * & * \end{pmatrix}$ for an $n \times n$ A. The way to do this is to to eliminate from the second row and multiplying the inverse from the left. Then do the same thing to the next column. It is easiest shown by an example.

Example 7 *Consider the matrix* $A = A_0 = \begin{pmatrix} 1 & 2 & 2 & 0 \\ 2 & 1 & 2 & 1 \\ 2 & 3 & 1 & 2 \\ 2 & 0 & 1 & 2 \end{pmatrix}$. *Do Gauss-elimination so that* $U_0A_0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 2 & 0 \\ 2 & 1 & 2 & 1 \\ 2 & 3 & 1 & 2 \\ 2 & 0 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 2 & 0 \\ 2 & 1 & 2 & 1 \\ 0 & 2 & -1 & 1 \\ 0 & -1 & -1 & 1 \end{pmatrix}$.
Then multiply the inverse

$$U_0A_0U_0^{-1} = \begin{pmatrix} 1 & 2 & 2 & 0 \\ 2 & 1 & 2 & 1 \\ 0 & 2 & -1 & 1 \\ 0 & -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 4 & 2 & 0 \\ 2 & 4 & 2 & 1 \\ 0 & 2 & -1 & 1 \\ 0 & -1 & -1 & 1 \end{pmatrix} = A_1.$$

We see now that

$$U_1A_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} 1 & 4 & 2 & 0 \\ 2 & 4 & 2 & 1 \\ 0 & 2 & -1 & 1 \\ 0 & -1 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 4 & 2 & 0 \\ 2 & 4 & 2 & 1 \\ 0 & 2 & -1 & 1 \\ 0 & 0 & -\frac{3}{2} & \frac{3}{2} \end{pmatrix}.$$

Multiply the inverse

$$U_1A_1U_1^{-1} = \begin{pmatrix} 1 & 4 & 2 & 0 \\ 2 & 4 & 2 & 1 \\ 0 & 2 & -1 & 1 \\ 0 & 0 & -\frac{3}{2} & \frac{3}{2} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 4 & 2 & 0 \\ 2 & 4 & \frac{3}{2} & 1 \\ 0 & 2 & -\frac{3}{2} & 1 \\ 0 & 0 & -\frac{9}{4} & \frac{3}{2} \end{pmatrix}$$

Set $U = U_0U_1$ *and we get that*

$$UAU^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -\frac{3}{2} & \frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 2 & 0 \\ 2 & 1 & 2 & 1 \\ 2 & 3 & 1 & 2 \\ 2 & 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & -\frac{1}{2} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 4 & 2 & 0 \\ 2 & 4 & \frac{3}{2} & 1 \\ 0 & 2 & -\frac{3}{2} & 1 \\ 0 & 0 & -\frac{9}{4} & \frac{3}{2} \end{pmatrix}$$

This method can be useful if you want to determinant the characteristic poly-

[2]More about this in: Matrix Computations third edition, Gene H. Golub,Charles F. Van Loan, The Johns Hopkins University press 1996 7.4

nomial of a matrix. Consider the matrix $H = \begin{pmatrix} h_{11} & h_{12} & \cdots & & h_{1n} \\ h_{21} & h_{22} & & & \\ 0 & h_{32} & & & \\ \vdots & \ddots & \ddots & & \\ 0 & \cdots & 0 & h_{n(n-1)} & h_{nn} \end{pmatrix}$ now if every $h_{j(j-1)} \neq 0$ and we have that $v = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ then the matrix $P = (v, Hv, H^2v, ..., H^{n-1}v)$ will be invertible(this is easy to check) and we can see that

$$P^{-1}HP = P^{-1}(Hv, H^2v, ..., H^nv) = \begin{pmatrix} 0 & 0 & \cdots & 0 & a_n \\ 1 & 0 & & & \vdots \\ 0 & 1 & & & \\ \vdots & \ddots & \ddots & & \\ 0 & \cdots & 0 & 1 & a_1 \end{pmatrix}.$$

We can after this calculation see that the characteristic polynomial of H is $s^n - a_1s^{n-1} - ... - a_n$ this can be verified by calculating $det(Is - H) = det(P)det(P^{-1}HP)det(P^{-1} = det(P^{-1}HP) =$

$$= \begin{vmatrix} s & 0 & \cdots & 0 & -a_n \\ -1 & s & & & \vdots \\ 0 & -1 & \ddots & & \\ \vdots & \ddots & \ddots & s & \\ 0 & \cdots & 0 & -1 & s - a_1 \end{vmatrix} = s^n - a_1s^{n-1} - ... - a_n$$

The last step follows from the definition of determinate. Finally note that if $h_{j(j-1)} = 0$ we can split computation of the characteristic polynomial into two smaller matrices

$$\begin{pmatrix} h_{11} & h_{12} & \cdots & & h_{1j} \\ h_{21} & h_{22} & & & \\ 0 & h_{32} & & & \\ \vdots & \ddots & \ddots & & \\ 0 & \cdots & 0 & h_{(j-1)(j-2)} & h_{(j-1)(j-1)} \end{pmatrix} \text{ and } \begin{pmatrix} h_{jj} & h_{j(j+1)} & \cdots & & h_{jn} \\ h_{(j+1)j} & h_{(j+1)(j+1)} & & & \\ 0 & h_{(j+2)(j+1)} & & & \\ \vdots & \ddots & \ddots & & \\ 0 & \cdots & 0 & h_{n(n-1)} & h_{nn} \end{pmatrix}.$$

We can now see the for any non singular matrix A we can decompose A into $P^{-1}HP$ where $H = \begin{pmatrix} C_1 & * & \cdots & * \\ 0 & C_2 & \ddots & \vdots \\ \vdots & & \ddots & * \\ 0 & \cdots & 0 & C_k \end{pmatrix}$ and $C_i = \begin{pmatrix} 0 & 0 & \cdots & 0 & * \\ 1 & 0 & & & \vdots \\ 0 & 1 & & & \\ \vdots & \ddots & \ddots & & \\ 0 & \cdots & 0 & 1 & * \end{pmatrix}$

from this we can always get the characteristic polynomial for A

Chapter 4

Eigenvalue problems

In this chapter I'm going to look at problems where I need eigenvalues of a matrix to solve the problem.

4.1 Minimal polynomial

A minimal polynomial[1] for a matrix $A \in \mathbb{R}^{n\times n}$ is the polynomial $p(s)$ with the lowest degree for which $p(A) = 0$. The first thing I'm going to show is how to minimize a singular $n \times n$ matrix.

Theorem 2 *If $A \in \mathbb{K}^{n\times n}$ is singular then A can be factorized to $KM = A$ where K and M are full rank matrices, non-square. Then the minimal polynomial is $p(x)x$ where $p(x)$ is the minimal polynomial of MK*

The proof of this is straight foreword $p(A)A = p(KM)KM = Kp(MK)M = K0M = 0$, and this is the minimal polynomial since there musts be at least one solution must be zero, also if there existed an other polynomial of lower rank such that $a(A) = 0$ then this polynomial must still have 0 as a solution and there for we can see that $a(A) = a'(A)A = Ka(MK)M$ and then a' must be the minimal polynomial of MK.
To make this more general I state the theorem:

Theorem 3 *The minimal polynomial of $A \in \mathbb{K}^{n\times n}$ where in this case $\mathbb{K}$ is algebraic closed and with distinct eigenvalues $\lambda_1, ..., \lambda_m$ is $\prod_{i=1}^{m}(x-\lambda_i)^{k_i}$. Here k_i is defined as $rank(A-I\lambda_i)^{k_i-1} > rank(A-I\lambda_i)^{k_i} = rank(A-I\lambda_i)^{k_i+1}$*

Note that $m \leq n$ in general. Assume that the characteristic polynomial of a matrix $A \in \mathbb{K}^{n\times n}$ is $a(s)$ and λ is an eigenvalue of A then we can factorize

[1] More of this in:A polynomial approach to linear Algebra,Paul A. Fuhrmann,Springer 2012, p93

$a(s)$ so that $a(s) = (s - \lambda)^p b(s)$ so that $b(\lambda) \neq 0$. Now we know that $0 = a(A) = (A - I\lambda)^p b(A)$. Rank factorize $b(A) = K_b M_b$. Thus $0 = a(A) = (A - I\lambda)^p K_b M_b$ and it is now clear that $a(A) = 0$ iff $(A - I\lambda)^p K_b = 0$ and since the row space of a matrix $B \in \mathbb{R}^{n\times n}$ is the same for B^k and B^{k+1} iff $rank(B^k) = rank(B^{k+1})$ we can draw the conclusion that the minimal i for which $(A-I\lambda)^p K_b = 0$ is $rank(A-I\lambda)^{i-1} > rank(A-I\lambda)^i = rank(A-I\lambda)^{i+1}$.

4.2 Jordan decomposition

Jordan decomposition may refer to many different things, but here we talk about Jordan canonical form. In general, a square complex matrix A is similar to a block diagonal matrix

$$J = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_p \end{pmatrix}$$

where each block J_i is a square matrix of the form

$$J_i = \begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}.$$

So there exists an invertible matrix P such that $P^{-1}AP = J$ is such that the only non-zero entries of J are on the diagonal and the superdiagonal. J is called the Jordan normal form of A. Each J_i is called a Jordan block of A. In a given Jordan block, every entry on the super-diagonal is 1.

What I am going to do here is to find the nonsingular matrix P. To this end we give a method using full rank decomposition of matrices to construct the so-called Jordan chains, whose definition will be made clear in a while.

Say that the matrix $A \in \mathbb{K}^{n\times n}$ has only one eigenvalue λ. Set $H = A - \lambda I_n$. We want to find vectors $v_1, ..., v_m$ such that $H^{i_k} v_k = 0$ and $H^{i_k - 1} v_k \neq 0$ and $P = (H^{i_1-1}v_1, ..., v_1, H^{i_2-1}v_2, ..., v_2, ..., H^{i_m-1}v_m, ..., v_m)$ is a invertible $n \times n$ matrix. Set i such that $rank(H^i) - rank(H^{i+1}) = 0$ and $rank(H^{i-1}) - rank(H^i) \neq 0$. Do the factorization described in 2.4 such that $H^i = K_i' M_i'$.
Conciser the lemma:
Set a matrix Y such that $\mathrm{Im}(Y) \subset (\mathrm{Ker}(M_k) \setminus \mathrm{Im}(K_k)) \cup \{0\}$ and Y has the

biggest possible rank (recall section 2.1.2). Then $rank(Y) = (rank(H^{k-1}) - rank(H^k)) - (rank(H^k) - rank(H^{k+1}))$
I will show this by referring to Theorem 1 We know that $rank(Y) = rank(\text{Ker}(M_k)) - rank(\text{Ker}(M_k) \cap \text{Im}(K_k)) = (rank(H^{k-1}) - rank(H^k)) - rank(\text{Ker}(M_k) \cap \text{Im}(K_k))$ and that $rank(H^{k+1}) = rank(M_k K_k) = rank(H^k) - rank(\text{Ker}(M_k) \cap \text{Im}(K_k)) \Leftrightarrow rank(\text{Ker}(M_k) \cap \text{Im}(K_k)) = rank(H^k) - rank(H^{k+1})$ and we get that $rank(Y) = (rank(H^{k-1}) - rank(H^k)) - (rank(H^k) - rank(H^{k+1}))$
This lemma is important since we want to find every biggest possible vector space within $(\text{Ker}(M_k) \setminus \text{Im}(K_k)) \cup \{0\}$. Now check:
$(rank(H^0) - rank(H^1)) - (rank(H^1) - rank(H^2) = p_1)$
$(rank(H^1) - rank(H^2)) - (rank(H^2) - rank(H^3) = p_2)$
$\vdots$
$(rank(H^{i-1}) - rank(H^i)) - (rank(H^i) - rank(H^{i+1}) = p_i)$
Find a basis $x_1, ..., x_m$for every $\text{Im}(Y) \subset (\text{Ker}(M_k) \setminus \text{Im}(K_k)) \cup \{0\}$ big as p_k. Now multiply with $M'^{\dagger}_{k-1}$ for an $x_t \in (\text{Ker}(M_k) \setminus \text{Im}(K_k)) \cup \{0\}$. Then we have that $v_k = M'^{\dagger}_{i_k-1} x_k$. I will first demonstrate that this is true by an example and then prove it.

Example 8 *Consider the matrix* $A = \begin{pmatrix} 1 & 2 & 0 & -1 \\ -1 & 3 & 0 & 0 \\ 1 & -3 & 2 & 2 \\ -1 & 1 & 0 & 2 \end{pmatrix}$ *This matrix has one eigenvalue 2.*

Set$H_1 = (A - 2I_4) = \begin{pmatrix} -1 & 2 & 0 & -1 \\ -1 & 1 & 0 & 0 \\ 1 & -3 & 0 & 2 \\ -1 & 1 & 0 & 0 \end{pmatrix}$. *Do full rank factorization* $H_1 =$

$$\begin{pmatrix} -1 & 2 & 0 & -1 \\ -1 & 1 & 0 & 0 \\ 1 & -3 & 0 & 2 \\ -1 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -2 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 2 & 0 & -1 \\ -1 & 1 & 0 & 0 \end{pmatrix} = K_1 M_1.$$

Set $H_2 = M_1 K_1 = \begin{pmatrix} -1 & 2 & 0 & -1 \\ -1 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -2 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix}$

Do full rank factorization $H_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} (-1, 1)$

set $H_3 = (1, -1) \begin{pmatrix} 1 \\ 1 \end{pmatrix} = (0)$. *We see that:*
$(rank(H^0) - rank(H^1)) - (rank(H^1) - rank(H^2) = 1)$
$(rank(H^1) - rank(H^2)) - (rank(H^2) - rank(H^3) = 0)$

$(rank(H^2) - rank(H^3)) - (rank(H^3) - rank(H^4) = 1)$
We Should now look for one vector in $(\mathrm{Ker}(M_1)\setminus \mathrm{Im}(K_1))$ *and one in* $(\mathrm{Ker}(M_3)\setminus \mathrm{Im}(K_3))$. *This can be done by using the methods from 2.1.1 and 2.1.2. and we find that* $\begin{pmatrix}0\\0\\1\\0\end{pmatrix} \in (\mathrm{Ker}(M_1)\setminus \mathrm{Im}(K_1))$ *and* $(2) \in (\mathrm{Ker}(M_3)\setminus \mathrm{Im}(K_3))$.

Now $M_2' = (-1,1)\begin{pmatrix}-1 & 2 & 0 & -1\\ -1 & 1 & 0 & 0\end{pmatrix} = \begin{pmatrix}0 & -1 & 0 & -1\end{pmatrix}$ *and* $M_2'^{\dagger} = \begin{pmatrix}0\\-1\\0\\1\end{pmatrix}\frac{1}{2}$

now we see that $M_2'^{\dagger}(2) = \begin{pmatrix}0\\-1\\0\\1\end{pmatrix}$. *Now we get the matrix*

$$P = (H^2\begin{pmatrix}0\\-1\\0\\1\end{pmatrix}, H^1\begin{pmatrix}0\\-1\\0\\1\end{pmatrix}, \begin{pmatrix}0\\-1\\0\\1\end{pmatrix}, \begin{pmatrix}0\\0\\1\\0\end{pmatrix}) = \begin{pmatrix}2 & -3 & 0 & 0\\ 2 & -1 & -1 & 0\\ -2 & 5 & 0 & 1\\ 2 & -1 & 1 & 0\end{pmatrix}$$

Now we can check that $P^{-1}AP = \begin{pmatrix}2 & 1 & 0 & 0\\ 0 & 2 & 1 & 0\\ 0 & 0 & 2 & 0\\ 0 & 0 & 0 & 2\end{pmatrix}$

To conclude this we can prove:

Theorem 4 *if* $A \in \mathbb{K}^{n\times n}$, $A^i = K_i'M_i'$ *for full rank matrices* K_k *and* M_k *and a matrix* Vx *such that* $M_{i-1}V = I$ *and* $x \in \mathrm{Ker}(M_i)$ *then* $(A^{i-1}Vx, ..., Vx)$ *has full rank.*

Proof
First of all, $M_k'V, 0 \leq k \leq i-1$ has full rank since $M_{i-1}V$ has full rank. And since $A^iVx = K_i'M_ix = 0$ we have that if $(A^{i-1}Vx, ..., Vx)$ dose not has full rank then there exist a non-zero vector y such that $0 = (A^{i-1}Vx, ..., Vx)y$ now if we multiply A then $A0 = 0 = (0, A^{i-1}Vx, ..., AVx)y$ this is true if $A^{i-1}Vx = 0$ and that is not true, $(A^{i-1}Vx, ..., AVx)$ does not have full rank. And if we multiply with A again we get $A0 = 0 = (0, 0, A^{i-1}Vx, ..., A^2Vx)y$ and this is true if$A^{i-1}Vx = 0$ and that is not true, or if $(A^{i-1}Vx, ..., AVx)$ does not have full rank. Now we can do this until $0 = (0, ..., 0, A^{i-1}Vx)y$ and since $A^{i-1}Vx \neq 0$ we are done.

Theorem 5 *if* $A \in \mathbb{K}^{n\times n}$, $x_1 \in \mathrm{Ker}(M_i) \setminus \mathrm{Im}(K_i)$, $x_2 \in \mathrm{Ker}(M_j) \setminus \mathrm{Im}(K_j)$

for $i > j$ and V_1, V_2 are matrices such that $M'_{i-1}V_1 = I$ and $M'_{j-1}V_2$ then $(A^{i-1}V_1x_1, ..., V_1x_1, A^{j-1}V_2x_2, ..., V_2x_2)$ has full rank.

Proof.
We know that $(A^{i-1}V_1x_1, ..., V_1x_1)$ and $(A^{j-1}V_2x_2, ..., V_2x_2)$ has full rank. Assume that there exist vectors x, y such that $0 = (A^{i-1}V_1x_1, ..., V_1x_1)x - (A^{j-1}V_2x_2, ..., V_2x_2)y$ (this is true iff $(A^{i-1}V_1x_1, ..., V_1x_1, A^{j-1}V_2x_2, ..., V_2x_2)$ not has full rank). Then $0 = A0 = (0, A^{i-1}V_1x_1, ..., AV_1x_1)x - (0, A^{j-1}V_2x_2, ..., AV_2x_2)y$ this is true if $\mathrm{Im}(A^{i-1}V_1x_1) = \mathrm{Im}(A^{j-1}V_2x_2)$ or if $(A^{i-1}V_1x_1, ..., AV_1x_1) \cap (A^{j-1}V_2x_2, ..., AV_2x_2) \neq \emptyset$, and we can do this in the same way as in the previous proof until $0 = (0, ..., 0, A^{i-1}V_1x_1, ..., A^{j-1}V_1x_1)x - (0, ..., 0)y$ and since $(A^{i-1}V_1x_1, ..., A^{j-1}V_1x_1)$ has full rank the only thing left to show is that $\mathrm{Im}(A^{i-1}V_1x_1) \neq \mathrm{Im}(A^{j-1}V_2x_2)$. and this is true since $\mathrm{Im}(A^{j-1}V_2x_2)) = \mathrm{Im}(K'_{j-1}x_2) \neq \mathrm{Im}(K'_{j-1}K_j...K_{i-1}x_1) = \mathrm{Im}(K'_{i-1}x_1) = \mathrm{Im}(A^iV_1x_1)$ and we are done.

Note: If $x'_2, x_2 \in \mathrm{Ker}(M_i)\backslash\mathrm{Im}(K_i)$ and (x'_2, x_2) is of full rank but $\mathrm{Im}((x'_2, x_2)) \cap \mathrm{Im}(K_j) \neq \emptyset$ then $(A^{i-1}V_1x_1, ..., V_1x_1, A^{j-1}V_2x_2, ..., V_2x_2, A^{j-1}V_2x'_2, ..., V_2x'_2)$ don't have full rank. Set $v = \begin{pmatrix} k_1 \\ \vdots \\ k_j \end{pmatrix}$ and a such that $(x_2, x'_2)\begin{pmatrix} 1 \\ a \end{pmatrix} \in \mathrm{Im}(K)$.
Then
$(A^{j-1}V_2x_2, ..., V_2x_2)v + (A^{j-1}V_2x'_2, ..., V_2x'_2)va =$
$A^{j-1}V_2x_2k_1 + ... + V_2x_2k_j + A^{j-1}V_2x'_2k_1a + ... + V_2x'_2k_ja =$
$A^{j-1}V_2(x_2+x'_2a)k_1 + ... + V_2(x_2+x'_2a)k_j = (A^{j-1}V_2(x_2+x'_2a), ..., V_2(x_2+x'_2a))v$
Since $(x_2 + x'_2a) \in \mathrm{Im}(K)$ the matrix,
$(A^{i-1}V_1x_1, ..., V_1x_1, A^{j-1}V_2x_2, ..., V_2x_2, A^{j-1}V_2x'_2, ..., V_2x'_2)$ doesn't have full rank.

4.3 Determination of the feedback matrix

In control theory, one of the important topics is stabilization of a system. Consider a linear system

$$\frac{d\mathbf{x}(t)}{dt} = A\mathbf{x}(t) + Bu(t).$$

where $\mathbf{x} : \mathbb{R}_+ \to \mathbb{R}^n$ is a state vector, and $u : \mathbb{R}_+ \to \mathbb{R}^m$ is an input or control variable, $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$.

In practice, we want to the system behaves as we wish, for example $x(t)$ generated by this system should go to 0, as $t \to \infty$. We know that ([4]) the eigenvalues of A play a essential role in this problem. If A has a "wrong" eigenvalue what can we do? In control theory we have freedom to choose u. A common strategy here is to use a feedback control law, i.e. choose $u(t) = Kx(t)$ by a proper choice of K with $K \in \mathbb{R}^{m\times n}$. Then we get a feedback system

$$\frac{d\mathbf{x}(t)}{dt} = (A + BK)\mathbf{x}(t).$$

In this section we shall give a method to find the matrix K, based on the machinery we have built up. We divide the problem into two cases (i) the so-called single-input case, i.e. B is just a column vector; and (ii) the so-called multi-input case, i.e. $m > 1$. For mathematical purposes I'm going to address the problem as.

$$\frac{d\mathbf{x}(t)}{dt} = (A - BK)\mathbf{x}(t).$$

4.3.1 Single-Input Case

I'm going to start with showing how you can do when $B := b \in \mathbb{R}^n$. For a matrix $A \in \mathbb{R}^{n\times n}$ with a characteristic polynomial $a(s) = s^n + a_1 s^{n-1} + ... + a_n$ and for the matrix $A - bk$ we have the characteristic polynomial $a_k(s) = s^n + \alpha_1 s^{n-1} + ... + \alpha_n$ which is often described by the location of the zeros.
Now we can show that $a_k(s) - a(s) = a(s)k(Is - A)^{-1}b \Leftrightarrow$
$a_k(s) = (k(Is - A)^{-1}b + 1)a(s) \Leftrightarrow det((sI - A + bk)(sI - A)^{-1}) =$
$= det(I + bk(sI - A)^{-1})$ this is clearly true.
We have $a_k(s) - a(s) = a(s)k(Is - A)^{-1}b$ and from this we can show that $\frac{1}{a(s)}[s^{n-1}I + s^{n-2}(A + Ia_1) + ...] = (sI - A)^{-1}$ and this can be shown as follows $\frac{1}{a(s)}[s^{n-1}I + s^{n-2}(A + Ia_1) + ...](sI - A) = \frac{1}{a(s)}(s^n I + s^{n-1}(A + Ia_1 - A) + s^{n-2}(A^2 + a_1 A + Ia_2 - (A^2 + a_1 A)) + ...) = \frac{1}{a(s)}a(s)I = I.$
Now we can simply see that:
$a_k(s) - a(s) = k[s^{n-1}I + s^{n-2}(A + Ia_1) + ...]b =$

$$k\begin{pmatrix} b & Ab & A^2b & \dots & A^{n-1}b \end{pmatrix} \begin{pmatrix} 1 & a_1 & \cdots & a_{n-1} \\ 0 & 1 & \cdots & a_{n-2} \\ 0 & 0 & \cdots & a_{n-3} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} s^{n-1} \\ s^{n-2} \\ \vdots \\ s \\ 1 \end{pmatrix} =$$

$$= \begin{pmatrix} \alpha_1 - a_1 & \alpha_2 - a_2 & \cdots & \alpha_n - a_n \end{pmatrix} \begin{pmatrix} s^{n-1} \\ s^{n-2} \\ \vdots \\ s \\ 1 \end{pmatrix} \Leftrightarrow$$

If $(b, Ab, ..., A^{n-1}b)$ is nonsigular then set:

$$k = \begin{pmatrix} \alpha_1 - a_1 & \alpha_2 - a_2 & \cdots & \alpha_n - a_n \end{pmatrix} \begin{pmatrix} 1 & a_1 & \cdots & a_{n-1} \\ 0 & 1 & \cdots & a_{n-2} \\ 0 & 0 & \cdots & a_{n-3} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}^{-1} \begin{pmatrix} b & Ab & A^2b & \ldots & A^{n-1}b \end{pmatrix}^{-1}$$

and we can see that we can choose any characteristic polynomial for $A - bk$ by chancing k. This is called controllable in control theory ([4]).

If $\begin{pmatrix} b & Ab & A^2b & \ldots & A^{n-1}b \end{pmatrix} \begin{pmatrix} 1 & a_1 & \cdots & a_{n-1} \\ 0 & 1 & \cdots & a_{n-2} \\ 0 & 0 & \cdots & a_{n-3} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = KM$ for full rank matrices M, K and if $\begin{pmatrix} \alpha_1 - a_1 & \alpha_2 - a_2 & \cdots & \alpha_n - a_n \end{pmatrix}^t \in \mathrm{Im}(M^t)$ then we can find an x such that $\begin{pmatrix} \alpha_1 - a_1 & \alpha_2 - a_2 & \cdots & \alpha_n - a_n \end{pmatrix} = xM$ and $k = x(K^tK)^{-1}K^t$.

Example 9 *If we have the matrix* $A = \begin{pmatrix} -4 & -4 & -7 & -4 \\ -3 & -2 & -4 & -1 \\ 6 & 5 & 10 & 5 \\ -3 & -1 & -4 & -2 \end{pmatrix}$ *and* $b = \begin{pmatrix} 2 \\ 2 \\ -2 \\ 1 \end{pmatrix}$ *and* A *has the characteristic polynomial* $a(s) = s^4 - 2s^3 - 3s^2 + 4s + 4 we$ *can build the matrix*

$$(b, Ab, A^2b, A^3b) \begin{pmatrix} 1 & -2 & -3 & 4 \\ 0 & 1 & -2 & -3 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 \end{pmatrix} =$$

$\begin{pmatrix} 2 & -6 & -5 \\ 2 & -3 & -2 \\ -2 & 7 & 9 \\ 1 & -2 & -3 \end{pmatrix} \begin{pmatrix} 1 & -2 & -3 & 0 \\ 0 & 1 & -2 & -3 \\ 0 & 0 & 1 & 1 \end{pmatrix} = KM$. *Now we see that* $(\alpha_1 + 2, \alpha_2 + 3, \alpha_3 - 4, \alpha_4 - 4)^T$ *must be in* $\mathrm{Im}(M^T)$. *Hence* $(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = x \begin{pmatrix} 1 & -2 & -3 & 0 \\ 0 & 1 & -2 & -3 \\ 0 & 0 & 1 & 1 \end{pmatrix} + (-2, -3, 4, 4)$ *for all* $x^T \in \mathbb{R}^3$

Theorem 6 *Assume $a(s)$ is the characteristic polynomial of the given matrix A and a desired polynomial $a_k(s)$ Then we have*
(i) If (A,b) is controllable then

$$k = \begin{pmatrix} \alpha_1 - a_1 & \alpha_2 - a_2 & \cdots & \alpha_n - a_n \end{pmatrix} \begin{pmatrix} 1 & a_1 & \cdots & a_{n-1} \\ 0 & 1 & \cdots & a_{n-2} \\ 0 & 0 & \cdots & a_{n-3} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}^{-1} \begin{pmatrix} b & Ab & A^2b & \ldots & A^{n-1}b \end{pmatrix}^{-1}$$

(ii) If (A,b) is not controllable but $(\alpha_1 - a_1, ...)^T \in \mathrm{Im}(M^t)$, then

$$k = x(K^T K)^{-1} K^T$$

for some $x \in \mathbb{R}^n$, where M and K are matrix factors of the full rank factorization of the matrix $\begin{pmatrix} b & Ab & A^2b & \ldots & A^{n-1}b \end{pmatrix} \begin{pmatrix} 1 & a_1 & \cdots & a_{n-1} \\ 0 & 1 & \cdots & a_{n-2} \\ 0 & 0 & \cdots & a_{n-3} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$.

Remark. In case (i) there are other methods to construct the matrix K see Sontag. Case (ii) corresponds the notion of stabilization of (A,b). So far I have not seen any precise construction of such K in literature.

4.3.2 Multi-Input Case

For the general problem I'm going to reduce the problem to the single input problem. Set $B = (b_1, ..., b_m)$ and $K = (k_1, ..., k_m)^T$ and for $rank(b_1, ..., A^{p-1}b_1) < rank(b_1, ..., A^p b_1) = rank(b_1, ..., A^{p+1}b_1)$ set k_2 such $A^p b_1 k_2^T \neq 0$ and $k_2 \in \mathrm{Ker}((b_1, ..., A^{p-1}b_1)^T)$ and then for
$rank(((b_1, ..., b_{i-1}), ..., A^{n-1}(b_1, ..., b_{i-1}), b_i, ..., A^{p_i - 1}b_i) <$

$rank(((b_1,...,b_{i-1}),...,A^{n-1}(b_1,...,b_{i-1}),b_i,...,A^{p_i}b_i) =$
$rank(((b_1,...,b_{i-1}),...,A^{n-1}(b_1,...,b_{i-1}),b_i,...,A^{p_i+1}b_i)$
set k_{i+1} such $A^{p_i}b_ik_{i+1}^T = -1$ and
$k_i \in \mathrm{Ker}((b_i,...,A^{p_i-1}b_i)^T) \oplus \mathrm{Ker}(((b_1,...,b_{i-1}),...,A^{n-1}(b_1,...,b_{i-1}))^T)$
for $i > 1$. If we can't pick such a k_{i+1} then set $k_{i+1} = 0$.
Set $A' = (A-(b_2,...,b_m)(k_2,...,k_m)^T)$ now we can see that $\mathrm{Im}(B, AB, ..., A^{n-1}B) = \mathrm{Im}(b_1, A'b_1, ..., A'^{n-1}b_1)$ this is true since we can see that $(b_1,...,A^{p_1}b_1) = (b_1,...,A'^{p_1}b_1)$ now consider $A'A^{p_1}b_1 = A^{p_1+1}b_1 + b_2$ and since $A^{p_1+1}b_1 \in \mathrm{Im}(b_1,...,A^{p_1}b_1)$ we don't have to care about that. Now consider $A'b_2 = Ab_2 + b_2k_2B_{b_2}$ and since $b_2k_2B_{b_2} \in (b_1,...,A'^{p_1+1}b_1)$ we don't have to care abut that ether. now we can see that $\mathrm{Im}((b_1,b_2),...,A^{n-1}(b_1,b_2)) = \mathrm{Im}(b_1, A'b_1, ..., A'^{n-1}b_1)$ and by similar arguments we can see that $\mathrm{Im}(B, AB, ..., A^{n-1}B) = \mathrm{Im}(b_1, A'b_1, ..., A'^{n-1}b_1)$
Now we have a single-input case that we can solve.

Example 10 *Consider* $A = \begin{pmatrix} 0 & 0 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & -2 \end{pmatrix}, B = \begin{pmatrix} 0 & -1 \\ 2 & 0 \\ 1 & 1 \end{pmatrix}$.
we can now see that $Ab_1 = \begin{pmatrix} 0 & 0 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & -2 \end{pmatrix} \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$ *and* $A^2b_1 = \begin{pmatrix} 0 & 0 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & -2 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$, *we can see that* $rank(b_1) < rank(b_1, Ab_1) = rank(b_1, Ab_1, A^2b_1)$. *I am going to pick* $k_2 = \begin{pmatrix} -\frac{1}{2} \\ 0 \\ 0 \end{pmatrix}$ *since* $Ab_1k_2^T = -1$ *and* $b_1k_2^T = 0$. *Set* $A' = \begin{pmatrix} 0 & 0 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & -2 \end{pmatrix} - \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} -\frac{1}{2} & 0 & 0 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} & 0 & 2 \\ 1 & 0 & 1 \\ \frac{1}{2} & 1 & -2 \end{pmatrix}$. *The characteristic polynomial of* A' *is* $s^3 + \frac{5}{2}s^2 - s - \frac{5}{2}$. *Now we can see that* $a_{k_1}(s) = s^3 + (k_1 \begin{pmatrix} 0 & 2 & -1 \\ 2 & 1 & 2 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & \frac{5}{2} & -1 \\ 0 & 1 & \frac{5}{2} \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} \frac{5}{2} & -1 & -\frac{5}{2} \end{pmatrix}) \begin{pmatrix} s^2 \\ s \\ 1 \end{pmatrix}$

Theorem 7 *consider the matrix* $n \times n$ $A + BK$ *where* $B \in \mathbb{R}^{n\times m}$ *and* $K \in \mathbb{R}^{m\times n}$ *where* A *and* B *are fixed. By choosing* K *the possible coefficients of the characteristic polynomial of* $A + BK$ *can be described by an linear equation.*

Note that I have not used any eigenvalues under this section.

Bibliography

[1] Linjär algebra med vektorgeometri, Anders Tengstrand, Anders Tengstrand och studentlitteratur 1994,2005

[2] A polynomial approach to linear Algebra, Paul A. Fuhrmann, Springer 2012

[3] Matrix Computations third edition, Gene H. Golub,Charles F. Van Loan, The Johns Hopkins University press 1996

[4] Mathematical control theory deterministic finite dimensional systems second edition, Eduardo D. Sontag, 1998 Springer-Verlag New York, Inc.