Tetration

Gottfried Helms - Univ Kassel   04 - 2008

Intro into Diagonalization

A short course

Abstract: a very elementary introduction into the idea of matrix-diagonalization is given.

Contents

1. *Diagnonalization – a short introduction*
   - 1.1. Intro: the diagonal matrix
   - 1.2. Matrix-operations under similarity transformations
   - 1.3. Diagonalization
2. *References*

1. Diagnonalization – a short introduction

1.1. Intro: the diagonal matrix

Let's start the inverse direction to that of the usual way of explaining this.

Assume a diagonal matrix ***D***: $\begin{bmatrix} a & . & . \\ . & b & . \\ . & . & c \end{bmatrix}$ D

Diagonal-matrices are the creatures with the nicest properties in the zoo of matrices:

You may add them,

$E = D + D = 2*D$ $\quad \begin{bmatrix} a & . & . \\ . & b & . \\ . & . & c \end{bmatrix} + \begin{bmatrix} a & . & . \\ . & b & . \\ . & . & c \end{bmatrix} = \begin{bmatrix} 2*a & . & . \\ . & 2*b & . \\ . & . & 2*c \end{bmatrix}$ D D 2*D

multiply them

$E = D * D = D^2$ $\quad \begin{bmatrix} a & . & . \\ . & b & . \\ . & . & c \end{bmatrix} * \begin{bmatrix} a & . & . \\ . & b & . \\ . & . & c \end{bmatrix} = \begin{bmatrix} a\hat{}2 & . & . \\ . & b\hat{}2 & . \\ . & . & c\hat{}2 \end{bmatrix}$ D D D*D

where we see the important property, that the computation of an element of the result for each of these operations ***involves only*** the elements of the source-matrices ***at the same position***.

The effect of this property is then, that we may express ***any*** function on a diagonal-matrix, which requires linear combination of these operations, say $D*D = D^2$, $exp(D) = I + D/1! + D^2/2! + ...$ etc

All this can be understood as application of these operators or functions to its diagonal-elements only, thus diagonal matrices allow any matrix-function which are allowed for its diagonal elements.

From this also the generalization to fractional powers of a diagonal-matrix is understandable: the root of the matrix is just the matrix of the root of its diagonal-elements. The meaningfulness of this can easily be checked: just try $\mathbf{E} = sqrt(\mathbf{D}) = diag(sqrt(a),sqrt(b),sqrt(c))$, compute $\mathbf{E}^2$ and see that $\mathbf{E}^2=\mathbf{D}$.

Same for any function, for another example $log(\mathbf{D})$ we have this

$E = log(D)$ $\quad log(\begin{bmatrix} a & . & . \\ . & b & . \\ . & . & c \end{bmatrix}) = \begin{bmatrix} log(a) & . & . \\ . & log(b) & . \\ . & . & log(c) \end{bmatrix}$ D E

So the meaningfulness of the construct "logarithm of a diagonal-matrix" lies in the fact, that logarithm can be expressed as powerseries, and each term can take the diagonal matrix as parameter giving a matrix-term, and all these matrix-terms can be summed in the obvious manner.

One most important entity is the (multiplicative) ***matrix-inverse***, as is computing $1/x$ from x in scalar arithmetic to complete the operation of multiplication:

$D^{-1} = diag(1/a,1/b,1/c)$ $\quad \begin{bmatrix} 1/a & . & . \\ . & 1/b & . \\ . & . & 1/c \end{bmatrix}$ D^-1

such that

$D*D^{-1} = D^{-1}*D = I$ *// is the identity*

Because of the simpliness of diagonal matrices, it is also, that their matrix-product commutes, so

$D*E = E*D$ // D,E diagonal

simply because the scalar multiplication (of their resp. diagonal-elements) commutes.

1.2. Matrix-operations under similarity transformations

Now assume another arbitray, but invertible matrix $\boldsymbol{W}$ and its inverse $\boldsymbol{W}^{-1}$ call it $\boldsymbol{WI}$ here:

Example:

$$W = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{bmatrix} \qquad WI = W^{-1} = \begin{bmatrix} 3 & -3 & 1 \\ -3 & 5 & -2 \\ 1 & -2 & 1 \end{bmatrix}$$

For the following we use numerical values for (a,b,c) in the diagonal of $\boldsymbol{D}$ to make simpler examples.

$$D = \begin{bmatrix} 1 & . & . \\ . & 2 & . \\ . & . & 3 \end{bmatrix}$$

Consider now the product $\boldsymbol{M = W*D*WI}$

$$M = W * D * WI \qquad \begin{bmatrix} 0 & 1 & 0 \\ 0 & -1 & 2 \\ 3 & -9 & 7 \end{bmatrix}$$

Because of the very useful properties of such a multiplication, where at the same time a matrix is pre- and postmultiplied by another matrix and its inverse, this operation has an own name: this is called a "***similarity transformation***"

The result $\boldsymbol{M}$ looks like a rather arbitrary matrix to us, as they occur in our daily work, and to apply matrix-functions to it seems to be a tedious task – for instance: well, we may compute the matrix-power $\boldsymbol{M}^2$ but how in the world could we find a matrix-root? $\boldsymbol{M}^{1/2}$?

Well, we go step-by-step...

We begin with the easiest operation, to demonstrate the use of our hidden knowledge about its composition.

$$M + M = W * D * WI + W * D * WI$$

Well, to compute the sum of two matrices is easy using element-by-element addition. But in the form of its composition we may do some algebraic operations, using associativy :

$$\begin{aligned} M + M &= W * D * WI + W * D * WI \\ &= W *(D * WI + D*WI) \\ &= W* ((D+D)*WI \\ &= W* (D+D)*WI \\ &= W* 2\ D * WI \\ &= 2\ W* D * WI \qquad \text{//since the scalar factor commutes with any matrix} \\ &= 2\ M \end{aligned}$$

More interesting is the matrix-product. Here we use the definition ***WI*W = I*** from the previous, and that we can omit any occuring identity-matrix in a matrix-product:

$$\begin{aligned} M * M &= W * D * WI * W * D * WI \\ &= W * D * (WI * W) * D * WI \\ &= W * D * (I) * D * WI \\ &= W * (D*D) * WI \\ &= W * D^2 * WI \qquad \text{// we compute this expression as most simple task} \\ &= M^2 \end{aligned}$$

While again the example would be easier to be computed directly instead via its components, we see immediately the perspectivic benefit: for higher powers it is the same, but reduced to the need of computation of the higher power just of the scalar-elements on the diagonal ***D***.

$$\begin{aligned} M^5 &= M * M * M * M * M \\ &= (W * D * WI) * (W * D * WI) * (W * D * WI) * (W * D * WI) * (W * D * WI) \\ &= W * D * (WI * W) * D * (WI * W) * D * (WI * W) * D * (WI * W) * D * WI \\ &= W * D * D * D * D * D * WI \\ &= W * D^5 * WI \qquad \text{// we compute this expression as most simple task} \end{aligned}$$

We see, that because of the construction ***WI*W = I*** the whole product collapses to a product of ***D***'s

If a function can be represented as a series, where – besides some scalar coefficients – only sums and powers of ***M*** occur, then the same result will be achieved, if we apply the function to the scalar diagonal-elements of ***D*** and finally compose

$$\begin{aligned} f(M) &= W * f(D) * WI \\ &= W * \operatorname{diag}(f(d_{0,0}), f(d_{1,1}), f(d_{2,2}), \ldots) * WI \end{aligned}$$

1.3. Diagonalization

So if we ***know*** the composition of ***M*** ahead, then everything is nice and we can use our knowledge to apply most functions to M as we would apply it to a scalar.

But what, if we come across the matrix ***M*** without that apriori-knowledge?

The we have to ***find*** this composition; we may say: we "decompose" ***M*** into three factors ***W***,***D***,***WI*** – if such is possible for a given ***M***.

The process of finding such a decomposition into factors for a given ***M*** is called "***diagonalization***" – because the primary goal is to find a diagonal matrix and two factors, which are multiplicative inverses of each other.

Because it has eminent use in mathematics (but also all over in engineereing and physics), and describes very important characteristics of an empirical matrix ***M***, any advanced algebra-software has the "eigensystem-solver" implemented, which tries to find such a diagonalization. The reader my look for "*eigenvalues*" (reflects the entries in ***D*** in this example), "*eigenvectors*" (reflect the column-vectors in ***W*** in this example), "*spectrum*" (refers to ***D*** as a diagonal matrix) or the more general approach "*Jordan-decomposition*" or "*jordan-form*" which give similar canonical forms for matrices, where the middle factor ***D*** cannot be diagonal.

The effect of the extinguishing ***WI*W*** in multiplicative formulae reminds also of the extinguishing of the exponent, if we have $b=t^{1/t}$ and repeat $b^t = t$ because $t^{1/t*t} = t^1 = t$ (but this just as a sidenote)

The numerical approach to diagonalization is not trivial, also some matrices cannot be represented as ***W*,*D*,*WI***-triple. There are some reformulations of the eigendecomposition, which are meant to induce practical ways to solve the problem of diagonalization, here is one example:

$$M = W * D * WI \quad => \quad M * W = W * D \quad \text{// by postmultiplication}$$

then for each column k of W we try to solve

$$M * W_k = d_{k,k} * W_k$$
$$M * W_k = (d_{k,k} * I) * W_k$$
$$(M - d_{kk} * I) * W_k = 0$$

and first find primarily valid $d_{k,k}$, known as the k'th eigenvalue and then to solve for $\boldsymbol{W_k}$.

This method is especially helpful with triangular matrices, since it can be shown, that -if they are diagonalizable- their diagonal entries are just their eigenvalues, and so for triangular matrices the diagonal ***D*** is known apriori.

So if ***D*** is not the identity-matrix ***I***, and all their entries differ (which is given with our "operator"-matrices in tetration) then the above formula leads to a recursive computation-scheme which gives exact entries (even for matrices where infinite size is assumed) for ***W*** and ***WI***.

The matrix-formula, which expresses the intention of "diagonalization" best is

$$M = W * D * W^{-1} \quad ==> \quad W^{-1} * M * W = D$$

which means, we look for $\boldsymbol{W^{-1}}$ and ***W*** , such that we can find the diagonal ***D*** for a given matrix ***M***.

After we have found this triple of matrices, we can apply functions like fractional powers, exponentials, logarithms and others in a meaningful and consistent way to matrices.

2. References

[Project-Index] *http://go.helms-net.de/math/tetdocs/index*

Gottfried Helms, 27.05.2008