LECTURE NOTES ON GENERALIZED EIGENVECTORS FOR SYSTEMS WITH REPEATED EIGENVALUES

We consider a matrix $A \in \mathbb{C}^{n\times n}$. The characteristic polynomial

$$P(\lambda) = |\lambda I - A|$$

admits in general p complex roots:

$$\lambda_1, \lambda_2, \ldots, \lambda_p$$

with $p \leq n$. Each of the root has a multiplicity that we denote k_i and $P(\lambda)$ can be decomposed as

$$P(\lambda) = \Pi_{i=1}^{p}(\lambda - \lambda_i)^{k_i}.$$

The sum of the multiplicity of all eigenvalues is equal to the degree of the polynomial, that is,

$$\sum_{i}^{p} k_i = n.$$

Let E_i be the subspace of eigenvectors associated to the eigenvalue λ_i, that is,

$$E_i = \{u \in \mathbb{C}^n \text{ such that } Au = \lambda_i u\}.$$

Theorem 1 (from linear algebra)**.** *The dimension of E_i is smaller than the multiplicity of λ_i, that is,*

$$\dim(E_i) \leq k_i.$$

If

$$\dim(E_i) = k_i \textbf{ for all } i \in \{1, \ldots, p\},$$

then we can find a basis of k_i independent eigenvectors for each λ_i, which we denote by

$$u_1^i, u_2^i, \ldots, u_{k_i}^i.$$

Since $\sum_{i=1}^{p} k_i = n$, we finally get **n linearly independent eigenvectors** (eigenvectors with distinct eigenvalues are automatically independent). Therefore the matrix A is diagonalizable and we can solve the system $\frac{dY}{dt} = AY$ by using the basis of eigenvectors. The general solution is given by

$$Y(t) = \sum_{i=1}^{p} e^{\lambda_i t}(a_{1,i}u_1^i + a_{2,i}u_2^i + \ldots + a_{k_i,i}u_{k_i}^i) \tag{1}$$

for any constant coefficients $a_{i,j}$.

example: We consider the matrix

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

Date: 14.03.2011.

The characteristic is $P(\lambda) = -(\lambda-2)(\lambda-1)^2$ and we have two eigenvalues, $\lambda_1 = 2$ (with multiplicity 1) and $\lambda_2 = 1$ (with multiplicity 2). We compute the eigenvectors for $\lambda_1 = 2$. We have to solve

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & -2 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0$$

It yields two independent relations (Hence, the dimension of $E_1 = n - 2 = 1$), namely

$$x = 0$$
$$y + z = 0$$

Thus, $u = (0, 1, -1)$ is an eigenvector for $\lambda_1 = 2$. We compute the eigenvectors for $\lambda_2 = 1$. We have to solve

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & -2 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0$$

It yields one independent relation, namely

$$2y + z = 0.$$

Hence, the dimension of E_2 is equal to $n - 1 = 2$. The two independent vectors

$$u_2^1 = (1, 0, 0)$$
$$u_2^2 = (0, 1, -2)$$

form a basis for E_2. Finally, the general solution of $\frac{dY}{dt} = AY$ is given by

$$Y(t) = a_1 e^{2t} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} + a_2 e^t \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + a_3 e^t \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}.$$

End of example.

If for some i, $\dim(E_i) < k_i$, then we cannot find k_i independent eigenvectors in E_i. We say that the eigenvalue λ_i is **incomplete**. In this case, we are not able to find n linearly independent eigenvectors and cannot get an expression like (1) for the solution of the ODE. We have to use **generalized eigenvectors**.

example: We consider

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

The characteristic polynomial is $P(\lambda) = (\lambda+2)^2$ and there is one eigenvalue $\lambda_1 = -2$ with multiplicity 2. We compute the eigenvectors. We have to solve

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0$$

It yields one independent relation, namely

$$y = 0$$

and therefore the dimension of E_1 is 1 and A is not diagonalizable. An eigenvector is given by $u_1 = (1, 0)$.

We know that $Y_1(t) = e^{\lambda_1 t} u_1$ is a solution. Let us look for solutions of the type

$$Y(t) = te^{\lambda t} u + e^{\lambda t} v$$

for two unknown vectors u and v different from zero. Such Y is solution if and only if

$$e^{\lambda t}u + \lambda t e^{\lambda t}u + \lambda e^{\lambda t}v = te^{\lambda t}Au + e^{\lambda t}Av$$

for all t. It implies that we must have

$$Au = \lambda u \tag{2}$$

$$Av = u + \lambda v. \tag{3}$$

The first equality implies (because we want $u \neq 0$) that u is an eigenvector and λ is an eigenvalue. We take $\lambda = \lambda_1$ and $u = u_1$. Let us compute v. We have to solve

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

which yields

$$y = 1$$

We take (for example) $v = (0, 1)$ and we have that

$$Y_2(t) = te^{\lambda_1 t}u_1 + e^{\lambda_1 t}v = te^{-2t}\begin{pmatrix} 1 \\ 0 \end{pmatrix} + e^{-2t}\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

is a solution. Since $Y_1(0)$ and $Y_2(0)$ are independent, a general solution is given by

$$\begin{aligned} Y(t) &= a_1 e^{-2t}\begin{pmatrix} 1 \\ 0 \end{pmatrix} + a_2\left(te^{-2t}\begin{pmatrix} 1 \\ 0 \end{pmatrix} + e^{-2t}\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) \\ &= \begin{pmatrix} a_1 e^{-2t} + a_2 t e^{-2t} \\ a_2 e^{-2t}. \end{pmatrix} \end{aligned}$$

for any constant a_1, a_2 in $\mathbb{R}$. End of example.

In the previous example, note that v satisfies

$$(A - \lambda I)^2 v = 0 \text{ and } (A - \lambda I)v \neq 0. \tag{4}$$

A vector which satisfies (4) is a generalized eigenvector. Let us give now the general definition of such vectors.

Definition 2. *For a given eigenvalue λ, the vector u is a **generalized eigenvector** of **rank** r if*

$$\begin{aligned} (A - \lambda I)^r u &= 0 \\ (A - \lambda I)^{r-1} u &\neq 0. \end{aligned}$$

Remark: An eigenvector is a generalized eigenvector of rank 1. Indeed, we have $(A - \lambda I)u = 0$ and $u \neq 0$.

Given an generalized eigenvector u of rank r, let us define the vectors $v_1, \ldots, v_r$ as follows

$$\begin{aligned} v_r &= (A - \lambda I)^0 u = u \\ v_{r-1} &= (A - \lambda I)^1 u \\ &\vdots \\ v_1 &= (A - \lambda)^{r-1} u \end{aligned} \tag{5}$$

Note that v_1 is an eigenvector as $v_1 \neq 0$ and $(A - \lambda I)v_1 = (A - \lambda)^r u = 0$. The vectors $v_1, \ldots, v_r$ form a **chain of generalized eigenvectors of length** r.

Definition 3. *Given an eigenvalue λ, we say that $v_1, v_2, \ldots, v_r$ form a* ***chain of generalized eigenvectors of length*** *r if $v_1 \neq 0$ and*

$$
\begin{aligned}
v_{r-1} &= (A - \lambda I)v_r \\
v_{r-2} &= (A - \lambda I)v_{r-1} \\
&\vdots \\
v_1 &= (A - \lambda I)v_2 \\
0 &= (A - \lambda I)v_1
\end{aligned}
\tag{6}
$$

Remark: Given a chain $\{v_i\}_{i=1}^r$, the first element, that is, v_1, is allways an eigenvalue. By using the definition (6), we get that

$$(A - \lambda I)^{i-1} v_i = v_1 \tag{7}$$

and therefore the element v_i is a generalized vector of rank i .

Theorem 4. *The vectors in a chain of generalized eigenvectors are linearly independent.*

Proof. We consider the linear combination

$$\sum_{i=1}^{r} a_i v_i = 0. \tag{8}$$

By using the definition (6), we get that

$$v_i = (A - \lambda I)^{r-i} v_r$$

so that (8) is equivalent to

$$\sum_{i=1}^{r} a_i (A - \lambda I)^{r-i} v_r = 0. \tag{9}$$

We want to prove that all the a_i are equal to zero. We are going to use the fact that

$$(A - \lambda I)^m u = 0 \text{ for all } m \geq r. \tag{10}$$

Indeed,

$$(A - \lambda I)^m v_r = (A - \lambda I)^{m-r} (A - \lambda I)^r v_r = (A - \lambda I)^{m-r} (A - \lambda I) v_1 = 0.$$

Now, we apply $(A - \lambda I)^{r-1}$ to (9) and get

$$\sum_{i=1}^{r} a_i (A - \lambda I)^{2r-i-1} v_r = 0. \tag{11}$$

Since $(A - \lambda I)^{2r-i-1} v_r = 0$ for $i \leq r - 1$, the equation (11) simplifies and we get

$$a_r (A - \lambda I)^{r-1} v_r = a_r v_1 = 0$$

Hence, $a_r = 0$ because $v_1 \neq 0$. Now, we know that $a_r = 0$ so that (9) rewrites as

$$\sum_{i=1}^{r-1} a_i (A - \lambda I)^{r-i} v_r = 0. \tag{12}$$

We apply $(A - \lambda I)^{r-2}$ to (12) and obtain that

$$\sum_{i=1}^{r-1} a_i (A - \lambda I)^{2r-i-2} v_r = a_{r-1} (A - \lambda I)^{r-1} v_r = a_{r-1} v_1 = 0.$$

because $(A-\lambda I)^{2r-i-2}v_r = 0$ for $i \leq r-2$. Therefore, $a_{r-1} = 0$. We proceed recursively with the same argument and prove that all the a_i are equal to zero so that the vectors v_i are linearly independent. □

A chain of generalized eigenvectors allow us to construct solutions of the system of ODE. Indeed, we have

Theorem 5. *Given a chain of generalized eigenvector of length r, we define*

$$\begin{aligned}
X_1(t) &= v_1 e^{\lambda t}\\
X_2(t) &= (tv_1 + v_2)e^{\lambda t}\\
X_3(t) &= \left(\frac{t^2}{2}v_1 + tv_2 + v_3\right)e^{\lambda t}\\
&\vdots\\
X_r(t) &= \left(\frac{t^{r-1}}{(r-1)!}v_1 + \ldots + \frac{t^2}{2}v_{r-2} + tv_{r-1} + v_r\right)e^{\lambda t}
\end{aligned}$$

The functions $\{X_i(t)\}_{i=1}^r$ form r linearly independent solutions of $\frac{dX}{dt} = AX$.

Proof. We have

$$X_j(t) = e^{\lambda t}\left(\sum_{i=1}^{j}\frac{t^{j-i}}{(j-i)!}v_i\right).$$

We use the convention that $v_0 = 0$ and, with this convention, we can check from (6) that

$$Av_i = v_{i-1} + \lambda v_i$$

for $i = \{1, \ldots, r\}$. We have, on one hand,

$$\dot{X}_j(t) = e^{\lambda t}\sum_{i=1}^{j-1}\frac{t^{j-i-1}}{(j-i-1)!}v_i + e^{\lambda t}\sum_{i=1}^{j}\lambda\frac{t^{j-i}}{(j-i)!}v_i$$

and, on the other hand,

$$\begin{aligned}
AX_j(t) &= e^{\lambda t}\sum_{i=1}^{j}\frac{t^{j-1}}{(j-i)!}Av_i\\
&= e^{\lambda t}\sum_{i=1}^{j}\frac{t^{j-i}}{(j-i)!}(v_{i-1} + \lambda v_i)\\
&= e^{\lambda t}\sum_{i=1}^{j}\frac{t^{j-i}}{(j-i)!}v_{i-1} + e^{\lambda t}\sum_{i=1}^{j}\lambda\frac{t^{j-i}}{(j-i)!}v_i\\
&= e^{\lambda t}\sum_{i=1}^{j-1}\frac{t^{j-i-1}}{(j-i-1)!}v_i + e^{\lambda t}\sum_{i=1}^{j}\lambda\frac{t^{j-i}}{(j-i)!}v_i.
\end{aligned}$$

Hence, X_i is a solution. To prove that $X_i(t)$ are independent, it is enough to prove that $X_i(0) = v_i$ are independent. This follows from Theorem 4. □

Conclusion: A chain of generalized eigenvectors of length r gives us r independent solutions.

It turns out that there exist enough chains of generalized eigenvectors to obtain a complete set of independent solutions. This is the content of the following theorem.

Theorem 6 (from linear algebra)**.** *For an eigenvalue λ of multiplicity k, there exist p chains, which we denote*

$$\begin{array}{cccc} v_1^1, & v_1^2, & \dots & ,v_1^p \\ v_2^1 & v_2^2 & & v_2^p \\ \vdots & \vdots & & \vdots \\ \vdots & v_{r_2}^2 & & v_{r_p}^p \\ v_{r_1}^1 & & & \end{array}$$

such that the $\{v_i^j\}$ are all linearly independent ***and*** *$\sum_i^p r_i = k$. Here r_i denotes the length of the i^{th} chain.*

Conclusion: For a general matrix (see first page), we have p eigenvalues, $\{\lambda_i\}_{i=1}^p$, where each of them has mutliplicity k_i. The sum of the multiplicities is equal to the dimension of the system n, that is, $\sum_{i=1}^p k_i = n$. For each eigenvalue λ_i, we compute k_i independent solutions by using Theorems 5 and 6. We finally obtain n **independent solutions** and find the general solution of the system of ODEs.

The following theorem is very usefull to determine if a set of chains consist of independent vectors.

Theorem 7 (from linear algebra)**.** *Given p chains, which we denote in the same way as in Theorem 6, the vectors $\{v_i^j\}$ are independent if and only if the*

$$v_1^1, \quad v_1^2, \quad \dots \quad ,v_1^p$$

are independent. This statement also holds if the chains correspond to different eigenvalues.

The proof of this theorem (in the case of the same eigenvalue) is of the same flavor as the proof of theorem 4.

Theorem 6 says that there exists a basis where the matrix can be rewritten as

$$D = \begin{pmatrix} \lambda_1 & 1 & & & & & & & \\ & \ddots & \ddots & & & & & & \\ & & \ddots & 1 & & & & & \\ & & & \lambda_1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & \lambda_q & 1 & & \\ & & & & & & \ddots & \ddots & \\ & & & & & & & \ddots & 1 \\ & & & & & & & & \lambda_q \end{pmatrix}$$

In this matrix, the only non zero entries are on the diagonal and the superdiagonal. The diagonal is made of the eignvalues while the superdiagonal contains either zero or one. For any matrix A, there exists therefore a matrix D of the form above and a change of base matrix P such that A writes

$$A = P^{-1}DP.$$

This decomposition of an arbitrary matrix in an *allmost diagonal* matrix of this type is called the **Jordan decomposition**.

The question is now: **How do we compute the p chains of generalized eigenvectors of Theorem 6**?

We use the following theorem of linear algebra.

Theorem 8 (from linear algebra)**.** *Given an eigenvalue λ of multiplicity k, let m be the dimension of the the subspace of eigenvectors. Then, for any generalized eigenvector u, we have*

$$(A - \lambda I)^{k-m+1}u = 0.$$

Basically, this theorem says that when we have a generalized eigenvector u of rank r (and therefore $(A - \lambda I)^r u = 0$ and $(A - \lambda I)^{r-1}u \neq 0$) then r cannot be larger than $k - m + 1$.

Now, we can present an algorithm to find the chains of independent generalized eigenvectors of Theorem 6 for a given eigenvalue λ.

First, we compute the eigenvectors and find the dimension m of the subspace of eigenvectors. We compute $(A - \lambda I)^{k-m+1}$.

Let $\mathcal{E}$ be a collection of independent chains that we are going to construct iteratively.

We start with a collection $\mathcal{E}$ that contains a basis of eigenvectors (they have been computed previously). Recall that an eigenvector is also a chain of length 1.

While the collection $\mathcal{E}$ does not contain a total number of k vectors, **do**

- Find a vector u satisfying $(A - \lambda I)^{k-m+1}u = 0$ and which is independent of the vectors that are contained in (the chains of) $\mathcal{E}$.
- Compute $(A-\lambda I)^j u$ until we find the largest j, that we denote r, such that
$$(A - \lambda I)^r u = 0 \text{ and } (A - \lambda I)^r u \neq 0.$$
Thus, u is a generalized vector of rank r.
- Construct the chain
$$\begin{aligned} v_r &= u \\ v_{r-1} &= (A - \lambda I)u \\ &\vdots \\ v_1 &= (A - \lambda I)^{r-1}u. \end{aligned}$$
- If there is a subset of chains in $\mathcal{E}$ which are not linearly independent with the chain $\{v_1, \dots, v_r\}$ (use Theorem 7 to test it) then remove among those chains the one with the smallest length otherwise do nothing. Add the chain $\{v_1, \dots, v_r\}$ to the collection $\mathcal{E}$.

end do

Note that the number of vectors in $\mathcal{E}$ strictly increases at each interation in the while-loop so that the loop will stop in a finite number of iteration.

example We consider

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

We compute the characteristic polynomial and find $P(\lambda) = (\lambda - 1)^3$. Hence, $\lambda_1 = 1$ is an eigenvalue of multiplicity 3. We compute the eigenvectors. We have to solve

$$\begin{pmatrix} 1 & 1 & 0 \\ -1 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0$$

We obtain one independent relation, namely,

$$x = -y$$

and therefore the dimension of E_1 is equal to $n - 1 = 2$. We have that

$$u_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \text{ and } u_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

are two independant eigenvectors. Our collection of chain $\mathcal{E}$ consist now of the eigenvalues u_1 and u_2, that is,

$$\mathcal{E} = \{\{u_1\}, \{u_2\}\}.$$

We compute and get $(A - \lambda_1 I)^2 = 0$. Hence, for any vector u, we have $(A - \lambda_1 I)^2 u = 0$. We choose (for example) $u = (1, 0, 0)$ so that u is linearly independent of u_1 and u_2. We get

$$(A - \lambda_1 I)v = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

and we get the chain

$$v_2 = u, \quad v_1 = (A - \lambda_1 I)u = u_1$$

We remove from $\mathcal{E}$ the chain $\{u_1\}$ as the vector which composes this chain is linearly dependent of the vectors of the chain $\{v_1, v_2\}$. We add the chain $\{v_1, v_2\}$. We end up with

$$\mathcal{E} = \{\{v_1, v_2\}, \{u_2\}\}.$$

We now have a basis $\{v_1, v_2, u_2\}$ of $\mathbb{R}^3$ whose vectors can be ordered in chains. By applying theorem 5, we get that

$$\begin{aligned} X(t) &= a_1 e^t u_2 + a_2 e^t v_1 + a_3 e^t (t v_1 + v_2) \\ &= a_1 e^t \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + a_2 e^t \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + a_3 e^t \begin{pmatrix} t+1 \\ -t \\ 0 \end{pmatrix} \\ &= e^t \begin{pmatrix} a_2 + a_3(t+1) \\ -a_2 - a_3 t \\ a_1 \end{pmatrix} \end{aligned}$$

is a general solution.