Matrices depending on parameters

(notes for M4A35)

Jeroen Lamb

22 October 2001

Introduction

The reduction to Jordan normal form (see appendix) is not a stable operation if we have multiple eigenvalues. Indeed, in the presence of multiple eigenvalues, an arbitrarily small perturbation may change the normal form completely. Actually, typically, a matrix does not have multiple eigenvalues, so this problem would not arise.

However, multiple eigenvalues may be unremovable under small perturbations when we are interested in a family of matrices depending on parameters, rather than in an individual matrix. In this case, although we can transform each individual matrix into Jordan normal form, both this normal form and the coordinate transformation leading to it will generally depend discontinuously on the parameter.

Consequently, we address the following problem:
What is the simplest form to which a family of matrices depending smootly on parameters can be reduced by a change of coordinates depending smoothly on the parameters?

We are mainly interested in this questions applied to real matrices $A \in gl(n, \mathbb{R}) \cong \mathbb{R}^{n^2}$, and real coordinate transformations $C \in Gl(n, \mathbb{R})$ (nonsingular matrices).

However, we will first consider the slightly easier context of complex matrices $A \in gl(n, \mathbb{C}) \cong \mathbb{C}^{n^2}$, and coordinate transformations $C \in Gl(n, \mathbb{C})$ (nonsingular matrices).

Versal deformations in $gl(n, \mathbb{C})$

Definition 1 (Equivalence) Let $A, B \in gl(n, \mathbb{C})$. We say that $A \cong B$ iff there exists a $C \in Gl(n, \mathbb{C})$such that $B = CAC^{-1}$.

This defines equivalence classes of matrices, partitioning $\mathbb{C}^{n^2}$ into manifolds that are orbits of the $Gl(n, \mathbb{C})$:

$$Orb(A) = \{B \in gl(n, \mathbb{C}) \mid B = CAC^{-1},\ C \in Gl(n, \mathbb{C})\}.$$

We consider a (real analytic) family of matrices

$$A : \mathbb{R}^\ell \to gl(n, \mathbb{C}) \cong \mathbb{C}^{n^2}.$$

We say that $A(\lambda)$ is a <u>deformation</u> of $A(0)$.

A deformation $A'$is <u>equivalent</u> to a deformation A if there exists a $C : \mathbb{R}^\ell \to Gl(n, \mathbb{C})$ with $C(0) = I$ such that

$$A'(\lambda) = C(\lambda)A(\lambda)C(\lambda)^{-1}.$$

Definition 2 The family *induced from A* by a mapping ϕ is the family

$$\phi^* A(\lambda) = A(\phi(\lambda)).$$

Definition 3 (Versal, universal, miniversal deformation) A defomation A of $A(0)$ is said to be *versal*if any deformation A' of $A(0)$ is equivalent to a deformation induced from A. A versal deformation is said to be *universal* if the induced mapping is determined uniquely be the deformation A'. A versal deformation is said to be *miniversal* if the dimension of the parameter space is the smallest possible for a versal deformation.

It is now our aim to construct versal deformations of matrices.

Let $A : \mathbb{R}^\ell \to \mathbb{C}^{n^2}$ be a deformation of $A(0) = A_0$ with parameter $\lambda \in \Lambda$, such that the mapping A is transversal $Orb(A_0)$ (ie A is not tangent to $Orb(A)$). We assume that the number of parameters of the deformation is minimal (ie equal to the codimension of $Orb(A_0)$ in $\mathbb{C}^{n^2}$ (dimension of the complement of $Orb(A_)$ in $\mathbb{C}^{n^2}$. Such a deformation is said to be minitransversal.

Lemma 4 *A minitransversal deformation A is miniversal.*

We will work towards proving this important Lemma.

Definition 5 The *centralizer* of a matrix $A \in gl(n, \mathbb{C})$ is the set of all matrices in $gl(n, \mathbb{C})$ commuting with A:

$$Z(A) = \{M \in gl(n, \mathbb{C}) \mid [M, A] = 0\},$$

where $[M, A] := MA - AM$.

$Z(A)$ is a linear subspace of $gl(n, \mathbb{C}) \cong \mathbb{C}^{n^2}$.

Consider the map $\psi : Gl(n, \mathbb{C}) \to gl(n, \mathbb{C})$, where $\psi(b) = bA_0b^{-1}$. Then, the derivative ψ' of ψ at $b = I$ is $\psi'(I)M = [M, A_0]$. Namely, $(I + \varepsilon M)A_0(I + \varepsilon M)^{-1} = A_0 + \varepsilon[M, A_0] + O(\varepsilon^2)$. Hence,

$$\psi'(I)M := \lim_{\varepsilon \to 0} \frac{\psi(I + \varepsilon M) - \psi(I)}{\varepsilon} = [M, A_0].$$

$Z(A_0)$ contains all matrices in the direction of which A_0 does not change. Consequently, A_0 changes in all other directions so that

$$\dim Orb(A_0) = \dim \mathbb{C}^{n^2} - \dim Z(A_0),$$

and hence, that

$$\dim Z(A_0) = \operatorname{codim} Orb(A_0).$$

In $\mathbb{C}^{n^2}$ we now introduce the scalar product

$$\langle A, B \rangle = Tr(AB^*),$$

where $Tr(M)$ is the trace (sum of all diagonal elements) of M, and $M^* = \bar{M}^T$. Verify that $\langle A, B \rangle$ is equal to the natural (Hermitian) scalar product in $\mathbb{C}^{n^2}$, taking the view that $A, B \in \mathbb{C}^{n^2}$ (verify this!!).

Lemma 6 *A vector B with origin at A_0 is perpendicular to $Orb(A_0)$ if and only if* $[B^*, A_0] = 0$.

Proof: As $\psi'(I)M = [M, A_0]$, vectors tangent to $Orb(A_0)$ are of the form $[M, A_0]$. Orthogonality of B to these vectors means that for all M, $\langle [M, A_0], B\rangle = 0$, ie

$$0 = Tr([M, A_0]B^*) = Tr(MA_0B^* - A_0MB^*) = Tr([A_0, B^*]M).$$

As the latter should be satisfied for all M, if follows that $[A_0, B^*] = 0$.

Now we prove the following lemma:

Lemma 7 *Let*

$$\Phi : P \times \mathbb{R}^\ell \to gl(n, \mathbb{C}) \cong \mathbb{C}^{n^2}, \quad \Phi(p, \lambda) = pA(\lambda)p^{-1},$$

where $P \subset Gl(n, \mathbb{C})$ transversal to $I + Z(A_0)$.

Then, in the neighbourhood of $(I, 0)$, Φ is a diffeomorphism on $(\mathbb{C}^{n^2}, A_0)$.

Proof: The derivative of Φ at $(I, 0)$ wrt p is $\psi'(I)$ and the derivative with respect to λ is $A'(0)$ (ie $dA/d\lambda(0)$). We claim that these operators map the space tangent to P at I and the space tangent to $\mathbb{R}^\ell$ at 0 onto transversal spaces of the same dimension: the tangent space to P at I is mapped to the tangent space of $Orb(A_0)$ at A_0, and the tangent space at $\mathbb{R}^\ell$ is mapped to a space transversal to it. Consequently, the derivative of Φ at $(I, 0)$ is an isomorphism between linear space of dimension n^2. By the inverse function theorem, Φ is a local diffeomorphism.

Proof of Lemma 4. We consider p and λas coordinates of the point $\Phi(p, \lambda)$. Let $\hat{A} : (M, 0) \to (\mathbb{C}^{n^2}, A_0)$be any deformation of A_0. Let μ be the parameter of the deformation. We define $\lambda = \phi(\mu)$ by $\phi(\mu) = \lambda(\hat{A}(\mu))$ and the set $B(\mu) = p(\hat{A}(\mu))$. Then $\hat{A}(\mu) = B(\mu)A(\phi(\mu))B^{-1}(\mu)$, which proves the versality of the deformation A.

The minimality of the dimension of the basis of this deformation is evident. ∎

Consequently, we can summarize the conclusion as follows:

Theorem 8 (Centralizer deformation [Arnold]) *Let $A_0 \in gl(n, \mathbb{C})$. Then the deformation*

$$\{A = A_0 + B \mid [B, A_0^*] = 0,\ B \in gl(n, \mathbb{C})\}$$

is miniversal. We can parametrize this deformation by a parameter whose dimension is equal to that of $Z(A_0^*)$.

Proof: Miniversality folows from Lemma 4. Note that

$$[B^*, A_0] = 0 \Leftrightarrow [B, A_0^*] = 0.$$

Homework 9

(1) State the inverse function theorem (used above). Show that it can be proven using the implicit function theorem!

(2) Construct miniversal deformation for the following complex matrices ($\alpha, \beta \in \mathbb{C}$, $\alpha \neq \beta$):

$$\begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}, \begin{pmatrix} \alpha & 1 \\ 0 & \alpha \end{pmatrix}, \begin{pmatrix} \alpha & 0 \\ 0 & \alpha \end{pmatrix}, \begin{pmatrix} \alpha & 1 & 0 \\ 0 & \alpha & 1 \\ 0 & 0 & \alpha \end{pmatrix}, \begin{pmatrix} \alpha & 1 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{pmatrix}.$$

Are the versal unfoldings you constructed also universal? (Show why (not).)

Versal deformations in $gl(n,\mathbb{R})$

The procedure is similar for matrices in $gl(n,\mathbb{R})$. One may construct a minitransversal unfolding, with the deformation $A_0 + B$, where $[A_0, B^T] = 0$. Then one can prove that this unfolding is miniversal.

Theorem 10 (Centralizer deformation [Arnold]) *Let $A_0 \in gl(n,\mathbb{R})$. Then the deformation*

$$\{A = A_0 + B \mid [B, A_0^T] = 0,\ B \in gl(n,\mathbb{R})\}$$

is miniversal. We can parametrize this deformation by a parameter whose dimension is equal to that of $Z(A_0^T) := \{B \in gl(n,\mathbb{R}) \mid [B, A_0^T] = 0\}$.

(Alternatively, the versal deformation of a real matrix can also be obtained from the versal deformations of complex matrices, by decomplexification.)

Homework 11 Prove Theorem 10. (Follow the procedure in the complex case step-by-step.)

Bifurcation diagrams

We now use the above theory to study changes of the Jordan normal form structure of real matrices as parameters are varied.

Because of the applications we have in mind, we are interested in describing parameter families of matrices where at some point the eigenvalues pass through the imaginary axis (in the complex plane). Namely, in case the eigenvalues are away from the imaginary axis, it can be shown that the vector field behaves very stably wrt changes of parameters.

We list some important miniversal unfoldings in the real case. We define here the "codimension" of a matrix A_0 with eigenvalues on the imaginary axis as the number of parameters needed to see a matrix with these particular eigenvalues on the imaginary axis persistently (ie we ignore the precise position of eigenvalues away from the imaginary axis).

- 1a. codimension 1: Jordan block with a zero eigenvalues 0
- 2a. codimension 2: Jordan block with pair of purely imaginary eigenvalues $\pm i\omega$.
  2b. codimension 2: Jordan block with two zero eigenvalues 0^2

How do we unfold? Use the centralizer unfolding.

$$1.\ (0+\lambda),\quad 2.a \begin{pmatrix} \lambda_1 & -(\omega+\lambda_2) \\ \omega+\lambda_2 & \lambda_1 \end{pmatrix},\quad 2.b \begin{pmatrix} \lambda_1 & 1 \\ \lambda_2 & \lambda_1 \end{pmatrix}.$$

The notion of codimension as used above is strictly the dimension of the miniversal deformation of the Jordan block(s) with eigenvalues on the imaginary axis. However, sometimes we have different notions, motivated by the application. For instance, in bifurcation theory we are interested when the eigenvalues of the matrices pass through the imaginary axis, but often it is relevant to know where only upto the extend whether this is at 0 or at $\pm i\omega$ for *some* ω. For the unfolding 2a, this means that the deformation with λ_2 is not so interesting or relevant. Hence, sometimes 2a is refered to as a codimension 1 bifurcation.

In the unfolding 2b, the eigenvalues are $\lambda = \lambda_1 \pm \sqrt{\lambda_2}$.

Homework 12

(1) Compare the deformation 2a with its complex counterpart and verify that the defomation could be derived (by decomplexification) from the complex miniversal deformation.

(2) Consider homework 9 (2). Construct miniversal unfolding for these matrices in $gl(n,\mathbb{R})$ when $\alpha, \beta \in \mathbb{R}$.

(3) Consider the same homework 9 (2) question and the miniversal deformations in $gl(n,\mathbb{C})$. What real deformations are found by decomplexification? Compare these deformations with the direct computation of miniversal unfoldings in real case.

A Jordan Normal form

Definition 13 (Characteristic polynomial) Let $A \in gl(n,\mathbb{C})$. Then, the *characteristic polynomial* of A is defined to be $p_A(\lambda) = \det(A - \lambda I)$.

Definition 14 (Eigenvalues) Let $A \in gl(n,\mathbb{C})$. λ is an eigenvalue of A iff there exists an $z \in \mathbb{C}^n \setminus \{0\}$ such that $Az = \lambda z$.

Lemma 15 *λ is an eigenvalue of A iff $p_A(\lambda) = 0$.*

We may write p_A in the following way:

$$p_A(\lambda) = \Pi_{i=1}^n (\lambda_i - \lambda).$$

Then,

$$E = \{\lambda_i \mid i = 1, \ldots, n\}$$

is the set of eigenvalues of A.

Definition 16 A complex *Jordan block* is an upper triangular matrix B whose diagonal elements $B_{j,j}$ are all identical, $B_{j,j+1} = 1$ for $j = 1, \ldots, n-1$, and all other elements are equal to 0.

There is an analogous result for real matrices:

Definition 17 A real *Jordan block* is identical to a corresponding complex Jordan block when the eigenvalue of the block is real, or identical to the decomplexification of the corresponding complex Jordan block when the eigenvalue of the block is not real.

Definition 18 A matrix A (in $gl(n,\mathbb{C})$ or $gl(n,\mathbb{R})$) is said to be in Jordan normal form if A is blockdiagonal and entirely made up of Jordan blocks.

Theorem 19 (Jordan normal form (real and complex)) *$Orb(A)$ contains a matrix in Jordan normal form, whose diagonal elements form the set E. The Jordan normal form of a matrix A is unique, upto ordering of the Jordan blocks.*

For proofs and more background, see for instance [2].

Homework 20

(1) Can one read the structure of the Jordan normal form of a matrix from its characteristic polynomial. Why (not) ?

* How can we compute the Jordan normal form of a matrix? Can we do it numerically, reliably? [try to work out some examples, have a look at some linear algebra books in the library etc....]

Background literature (suggestions):

[1] V.I. Arnold. *Geometric methods in the theory of ordinary differential equations*, Grundlehren der mathematischen Wissenschaften **250**, Springer, Berlin, 1983 (chapter 6, §30)
[2] K. Nomizu. *Fundamentals of linear algebra*, McGraw-Hill, New York, 1966 (chapter 7,8: on Jordan Normal Form).

Some homework answers:

Homework 9 :

(1) **Theorem 21 (Inverse function theorem)** *Let $f : E \to \mathbb{C}4^n$, C^k, where $E \subset \mathbb{C}^n$, with $f'(a)$ invertible for some $a \in E$. Then there exist open sets $U, V \subset \mathbb{C}^n$ such that $a \in U$, $f(a) \in V$, and f is one-to-one on U, $f(U) = V$, and f^{-1} is C^k.*

Proof: In order to use the implicit function theorem, we write

$$g(x, \lambda) = f(a + x) - (b + \lambda).$$

Hence g satisfies $g(0,0) = 0$, $dg/d\lambda(0,0) = 1$ and $dg/dx(0,0) = f'(a)$ both invertible. Hence by the implicit function theorem for all $\lambda \in \mathbb{C}^n$ near 0 there exists a unique $x(\lambda)$ with $x(0) = 0$ such that $g(x(\lambda), \lambda) = 0$. At the same time, for all x near 0 there exists a unique $\lambda(x)$ with $\lambda(0) = 0$, so that $g(x, \lambda(x)) = 0$. We thus have $f(a + x) = b + \lambda(x)$ and $f(a + x(\lambda)) = b + \lambda$, where $x(\lambda)$ and $\lambda(x)$ are unique. Hence these functions must refer to the same set, so that $x(\lambda(x)) = x$ and $\lambda(x(\lambda)) = \lambda$. In other words f is locally one-to-one.

(2) Miniversal deformations: ($\lambda_i \in \mathbb{C}$)

- $\begin{pmatrix} \alpha + \lambda_1 & 0 \\ 0 & \beta + \lambda_2 \end{pmatrix}$: <u>universal</u> since any deformation A' maps uniquely to this miniversal deformation by the values of the (distinct) eigenvalues.
- $\begin{pmatrix} \alpha + \lambda_1 & 1 \\ \lambda_2 & \alpha + \lambda_1 \end{pmatrix}$: <u>universal</u> since eigenvalues are $-(\alpha + \lambda_1) \pm \sqrt{\lambda_2}$, $1:1$ function of λ_1, λ_2.
- $\begin{pmatrix} \alpha + \lambda_1 & \lambda_2 \\ \lambda_3 & \alpha + \lambda_4 \end{pmatrix}$: <u>not universal</u> since eigenvalues are not $1:1$function of $\lambda_1, \ldots, \lambda_4$ so that we can construct a deformation with a nonunique representative in terms of $\lambda_1, \ldots, \lambda_4$.

- $\begin{pmatrix} \alpha+\lambda_1 & 1 & 0 \\ \lambda_2 & \alpha+\lambda_1 & 1 \\ \lambda_3 & \lambda_2 & \alpha+\lambda_1 \end{pmatrix}$: appears <u>universal</u> as eigenvalues appear unique function of $\lambda_1, \lambda_2, \lambda_3$.
- $\begin{pmatrix} \alpha+\lambda_1 & 1 & 0 \\ \lambda_2 & \alpha+\lambda_1 & \lambda_3 \\ \lambda_4 & 0 & \alpha+\lambda_5 \end{pmatrix}$: <u>not universal</u>.

Homework 12 :

(1) 2a in complex form is $\alpha \in \mathbb{C}$ with $\alpha = i\omega$. The miniversal deformation of α is $\alpha + \lambda$. In real form (decomplexified), we write $\alpha = a + ib = a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, so that

$$\alpha + \lambda = (\lambda_1 + i(\omega + \lambda_2)) = \begin{pmatrix} \lambda_1 & -(\omega+\lambda_2) \\ \omega+\lambda_2 & \lambda_1 \end{pmatrix}.$$

(2) Same deformations, only now $\lambda_i \in \mathbb{R}$.

(3) Just some examples to illustrate the procedure:

In the first matrix, suppose $\alpha \in \mathbb{R}$ and $\beta = a + ib \in \mathbb{C}$. Then by decomplexification of the complex miniversal deformation we obtain

$$\begin{pmatrix} \alpha+\mu_1 & 0 \\ 0 & \beta+\mu_2 \end{pmatrix}. \quad \to \quad \begin{pmatrix} \alpha+\lambda_1 & 0 & 0 \\ 0 & a+\lambda_2 & -(b+\lambda_3) \\ 0 & b+\lambda_3 & a+\lambda_2 \end{pmatrix}.$$

This miniversal deformation is also obtained starting from the real matrix $\begin{pmatrix} \alpha & 0 & 0 \\ 0 & a & -b \\ 0 & b & a \end{pmatrix}$.

In the second matrix, suppose $\alpha = a + ib \in \mathbb{C}$. Then by decomplexification of the complex miniversal deformation we obtain (with $\mu_1 = \lambda_1 + i\lambda_2$ and $\mu_2 = \lambda_3 + i\lambda_4$)

$$\begin{pmatrix} \alpha+\mu_1 & 1 \\ \mu_2 & \alpha+\mu_1 \end{pmatrix}. \quad \to \quad \begin{pmatrix} a+\lambda_1 & -(b+\lambda_2) & 1 & 0 \\ b+\lambda_2 & a+\lambda_1 & 0 & 1 \\ \lambda_3 & -\lambda_4 & a+\lambda_1 & -(b+\lambda_2) \\ \lambda_4 & \lambda_3 & b+\lambda_2 & a+\lambda_1 \end{pmatrix}.$$

This miniversal deformation can also obtained starting from the real matrix $\begin{pmatrix} a & -b & 1 & 0 \\ b & a & 0 & 1 \\ 0 & 0 & a & -b \\ 0 & 0 & b & a \end{pmatrix}$.