Linearity

Linear ODEs form an important class of ODEs. They are characterized by the fact that the vector field $f : \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}^m$ is *linear* (at constant values of the parameters), that is for all $x, y \in \mathbb{R}^m$ and $a, b \in \mathbb{R}$

$$f(ax + by, \lambda) = af(x, \lambda) + bf(y, \lambda).$$

This implies that the vector field may be represented by an $m \times m$ matrix $A(\lambda)$, so that

$$\frac{dx}{dt} = A(\lambda)x.$$

Solutions form linear vector space

Proposition

If $x(t)$ and $y(t)$ are solutions of a linear ODE, then $ax(t) + by(t)$, with $a, b \in \mathbb{R}$, is also a solution.

Proof.

If

$$\frac{dx(t)}{dt} = f(x(t)) \text{ and } \frac{dy(t)}{dt} = f(y(t))$$

then

$$\frac{d(ax(t) + by(t))}{dt} = af(x(t)) + bf(y(t)) = f(ax(t) + by(t)).$$

□

Flows of linear ODEs are linear

Proposition

The flow of an ODE is linear if and only if the corresponding vector field is linear.

Proof.

First, if the flow Φ^t is linear, then by the fact that

$$\lim_{t \to 0} \frac{\Phi^t(x) - x}{t} = \frac{d}{dt}\Phi^t|_{t=0}(x) =: f(x)$$

it follows that f must be linear as well.
From the previous proposition, we see that if f is linear and we choose some initial condition $x + y$, then

$$\Phi^t(x+y) = x(t) + y(t) = \Phi^t(x) + \Phi^t(y).$$

□

We consider the linear autonomous ODE

$$\frac{dx}{dt} = Ax, \tag{3}$$

where $x \in \mathbb{R}^m$ and $A \in gl(m, \mathbb{R})$ (an $m \times m$ matrix with real entries).
In the case that $m = 1$, when $A \in \mathbb{R}$, the flow map Φ^t of the linear ODE is equal to (multiplication by) $\exp(At)$.

Exponential

It turns out that we can write the flow map of a general linear autonomous ODE in this same form, if we define the exponential operator as

Definition (Exponential of matrix)

Let $L \in gl(m, \mathbb{R})$, then

$$\exp(L) := \sum_{k=0}^{\infty} \frac{L^k}{k!}$$

You should recognize the familiar Taylor expansion of the exponential function in the case that $m = 1$. In the formula, L^k represents the matrix product of k consecutive matrices L.

On $\exp(L)$

It follows that $\exp(L)$ is also a linear map (which can be represented by an $m \times m$ matrix). Moreover, it is invertible, with inverse $\exp(-L)$.

Technically, one may worry about the fact whether the coefficients of the matrix $\exp(L)$ are indeed all smaller than infinity. But it is in fact not too hard to find an upper bound for any of the entries of $\exp(L)$ given that the entries of the matrix L are finite. (Exercise)

Theorem

The linear autonomous ODE with initial condition $x(0) = y$ *has the unique solution* $x(t) = \exp(At)y$.

Proof.

We use the definition of the derivative:

$$\frac{d}{dt}\exp(tA) = \lim_{h\to 0}\frac{\exp((t+h)A) - \exp(tA)}{h}$$
$$= \lim_{h\to 0}\frac{\exp(hA) - I}{h}\exp(tA) = \lim_{h\to 0}\left(\frac{1}{h}\sum_{n=1}^{\infty}\frac{(hA)^n}{n!}\right)\exp(tA)$$
$$= \lim_{h\to 0}\left(\frac{hA}{h} + \frac{1}{h}\sum_{n=2}^{\infty}\frac{(hA)^n}{n!}\right)\exp(tA)$$
$$= \lim_{h\to 0}\left(A + h\sum_{j=0}^{\infty}h^j\frac{A^{j+2}}{(j+2)!}\right)\exp(tA) = A\exp(tA).$$

Proof.

To show that the solution is unique, suppose that $x(t)$ is a solution. We calculate (by using the chain rule)

$$\begin{aligned} \frac{d}{dt}(\exp(-tA)x(t)) &= \left(\frac{d}{dt}\exp(-tA)\right)x(t)) + \exp(-tA)\frac{dx(t)}{dt} \\ &= -A\exp(-At)x(t) + \exp(-tA)Ax(t) = 0. \end{aligned}$$

Hence, $x(t) = \exp(tA)\mathbf{c}$ where $\mathbf{c} \in \mathbb{R}^m$ is some constant which is entirely determined to be equal to y by the boundary condition $x(0) = y$, and the solution $x(t)$ is unique. □

It is easy to extend the result to the case of initial condition $x(\tau) = y$, yielding $x(t) = \exp(A(t-\tau))y$.

Corollary (Existence and uniqueness)

For the linear autonomous ODE (3), through each point y in the phase space passes exactly one solution at time $t = \tau$, namely $x(t) = \exp((t - \tau)A)y$.

Suppose now that A has two distinct eigenvalues $\mu_1 \neq \mu_2$. Then it follows that the corresponding eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$ must be linearly independent, i.e. $\mathbf{v}_1 \neq a\mathbf{v}_2$ for all $a \in \mathbb{R}$. In other words, $\mathbf{v}_1$ and $\mathbf{v}_2$ span the plane $\mathbb{R}^2$. As a consequence, we may write any point as a linear combination of the two eigenvectors

$$x = a\mathbf{v}_1 + b\mathbf{v}_2, \tag{4}$$

and by linearity of the flow map, and knowing the flow in the directions of the eigenvectors, we obtain

$$\Phi^t(x) = \Phi^t(a\mathbf{v}_1 + b\mathbf{v}_2) = a\Phi^t(\mathbf{v}_1) + b\Phi^t(\mathbf{v}_2) = ae^{\lambda_1}\mathbf{v}_1 + be^{\lambda_2}\mathbf{v}_2. \tag{5}$$

In order to find the matrix expression for Φ^t we now need to find expressions for the coefficients a and b in (4). We do this by taking the (standard matrix) product on both sides with vectors $(v_i^\perp)^\top$, where $v_i^\perp$ is perpendicular to the eigenvector v_i, for $i = 1, 2$:

$$a = \frac{(\mathbf{v}_2^\perp)^\top x}{(\mathbf{v}_2^\perp)^\top \mathbf{v}_1}, \quad b = \frac{(\mathbf{v}_1^\perp)^\top x}{(\mathbf{v}_1^\perp)^\top \mathbf{v}_2}.$$

On a more abstract level, we may think of the vectors $a\mathbf{v}_1$ and $b\mathbf{v}_2$ being obtained from the vector x by *projection*:

$$P_1 x = a\mathbf{v}_1 = \mathbf{v}_1((\mathbf{v}_2^\perp)^\top \mathbf{v}_1)^{-1}(\mathbf{v}_2^\perp)^\top x,$$

$$P_2 x = b\mathbf{v}_2 = \mathbf{v}_2((\mathbf{v}_1^\perp \mathbf{v}_2)^\top)^{-1}(\mathbf{v}_1^\perp)^\top x.$$

Note that projections P are linear maps that satisfy the property that $P^2 = P$. As they are linear, we can represent them by matrices:

$$P_1 = \mathbf{v}_1((\mathbf{v}_2^{\perp})^{\top}\mathbf{v}_1)^{-1}(\mathbf{v}_2^{\perp})^{\top}, \quad P_2 = \mathbf{v}_2((\mathbf{v}_1^{\perp})^{\top}\mathbf{v}_2)^{-1}(\mathbf{v}_1^{\perp})^{\top}.$$

So in terms of these expressions, we can rewrite (5) as

$$\Phi^t = e^{\lambda_1 t}P_1 + e^{\lambda_2 t}P_2. \tag{6}$$

Example

Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$, then the eigenvalues are equal to $\lambda_1 = 1$ and $\lambda_2 = 2$ with corresponding eigenvectors $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Any vector in $\mathbb{R}^2$ can be decomposed into a linear combination of these eigenvectors as follows:

$$\begin{pmatrix} x \\ y \end{pmatrix} = (x-y)\begin{pmatrix} 1 \\ 0 \end{pmatrix} + y\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

$$\Rightarrow \quad \exp(At)\begin{pmatrix} x \\ y \end{pmatrix} = (x-y)e^t\begin{pmatrix} 1 \\ 0 \end{pmatrix} + ye^{2t}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$
$$= \begin{pmatrix} (x-y)e^t + ye^{2t} \\ ye^{2t} \end{pmatrix} = \begin{pmatrix} e^t & e^{2t} - e^t \\ 0 & e^{2t} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix},$$

Example (continued)

so that

$$\exp(At) = e^t \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} + e^{2t} \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

We now construct the projections using the formulas we obtained above. Choosing the perpendicular vectors (for instance) as $\mathbf{v}_1^\perp = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $\mathbf{v}_2^\perp = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, we obtain

$$P_1 = \mathbf{v}_1((\mathbf{v}_2^\perp)^\top \mathbf{v}_1)^{-1}(\mathbf{v}_2^\perp)^\top = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix},$$

$$P_2 = \mathbf{v}_2((\mathbf{v}_1^\perp)^\top \mathbf{v}_2)^{-1}(\mathbf{v}_1^\perp)^\top = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

We note that the above procedure can also be carried out in the higher dimensional case. In the m-dimensional case, if A has m distinct real eigenvalues $\lambda_1, \dots, \lambda_m$ then

$$\exp(At) = \sum_{k=1}^{m} e^{\lambda_k t} P_k,$$

where P_k is the projection to the one-dimensional subspace spanned by the eigenvector for eigenvalue λ_k with kernel equal to the $(m-1)$-dimensional subspace spanned by the collection of all other eigenvectors. Projections P are characterized by the following properties:

- P is linear
- $\mathrm{range}(P)$ (where to project to)
- $\ker(P)$ (what should be projected away)
- $P^2 = P$

Proposition (Projection formula)

A formula for the $m \times m$ matrix representing the projection with a given range and kernel can be found as follows. Let the vectors $\mathbf{u}_1, \dots, \mathbf{u}_k$ form a basis for the range of the projection, and assemble these vectors in the $m \times k$ matrix A. The range and the kernel are complementary spaces, so the kernel has dimension $m - k$. It follows that the orthogonal complement of the kernel has dimension k. Let $\mathbf{w}_1, \dots, \mathbf{w}_k$ form a basis for the orthogonal complement of the kernel of the projection, and assemble these vectors in the matrix B. Then the projection to the range is defined by

$$P = A(B^\top A)^{-1} B^\top.$$

It is readily verified that indeed the formulas we obtained for the projections in the two-dimensional case are special cases of this general result.

The eigenvalues of a matrix with real entries need not be real, but in case we have an eigenvalue $\lambda = \alpha + i\beta$ (with $\alpha, \beta \in \mathbb{R}$). Namely, eigenvalues λ of a matrix $A \in gl(m, \mathbb{R})$ are roots of the characteristic polynomial, that is, they satisfy

$$\det(A - \lambda I) = 0. \tag{7}$$

As A has only real entries, it follows (by taking the complex conjugate of the entire equation) that if λ satisfies (7) then so does its complex conjugate $\overline{\lambda}$.
The eigenvectors of a real matrix for complex eigenvalues are never in $\mathbb{R}^m$. This can be seen from taking the complex conjugate of the eigenvalue equation $A\mathbf{v} = \lambda\mathbf{v}$, from which it follows that if $\mathbf{v}$ is an eigenvector for eigenvalue λ then $\overline{\mathbf{v}}$ is an eigenvector for eigenvalue $\overline{\lambda}$.

Hence for each pair of complex conjugate eigenvalues $\lambda, \bar{\lambda}$ we have a complex conjugate pair of eigenvectors $\mathbf{v}, \overline{\mathbf{v}} \in \mathbb{C}^m$. Here we use the natural identification of $\mathbb{R}^m$ as a subspace of $\mathbb{C}^m$. Although practical, it is maybe not so elegant to go outside our "real" phase space $\mathbb{R}^m$.

In analogy to the case of real eigenvalues, we are interested in identifying an invariant subspace of $\mathbb{R}^m$ which is associated with the complex eigenvalues. The complex eigenvectors $\mathbf{v}, \overline{\mathbf{v}}$ span a complex vector space whose real subspace is given by

$$E = \mathrm{span}_{\mathbb{C}}(\mathbf{v}, \overline{\mathbf{v}}) \cap \mathbb{R}^2 = \mathrm{span}_{\mathbb{R}}(\mathbf{v} + \overline{\mathbf{v}}, i(\mathbf{v} - \overline{\mathbf{v}})).$$

$$E = \operatorname{span}_{\mathbb{C}}(\mathbf{v}, \overline{\mathbf{v}}) \cap \mathbb{R}^2 = \operatorname{span}_{\mathbb{R}}(\mathbf{v} + \overline{\mathbf{v}}, i(\mathbf{v} - \overline{\mathbf{v}})).$$

It is easily verified that indeed E is an A-invariant subspace: let $\lambda = \alpha + i\beta$ and $\mathbf{v} = x + iy$, then

$$\begin{aligned}
A(a(\mathbf{v} + \overline{\mathbf{v}}) + bi(\mathbf{v} - \overline{\mathbf{v}})) &= a(A\mathbf{v} + A\overline{\mathbf{v}}) + ib(A\mathbf{v} - A\overline{\mathbf{v}}) \\
&= a(\lambda\mathbf{v} + \overline{\lambda}\overline{\mathbf{v}}) + ib(\lambda\mathbf{v} - \overline{\lambda}\overline{\mathbf{v}}) \\
&= 2a(\alpha x - \beta y) - 2b(\alpha y + \beta x) \\
&= (2a\alpha - 2b\beta)x - (2a\beta + 2b\alpha)y \\
&= (a\alpha - b\beta)(\mathbf{v} + \overline{\mathbf{v}}) + (a\beta + b\alpha)i(\mathbf{v} - \overline{\mathbf{v}}).
\end{aligned}$$

With the vectors $\frac{\mathbf{v}+\bar{\mathbf{v}}}{|\mathbf{v}+\bar{\mathbf{v}}|}$ and $\frac{i(\mathbf{v}-\bar{\mathbf{v}})}{|i(\mathbf{v}-\bar{\mathbf{v}})|}$ as the (orthonormal) basis of E, let $A \in gl(2, \mathbb{R})$ has two complex conjugate eigenvalues. Then according to the above formula, with the above mentioned basis, it has he form

$$A = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}. \tag{8}$$

We note that the eigenvalues of this matrix are indeed $\alpha \pm i\beta$, as required, and the exponential can be computed

$$\exp(At) = e^{\alpha t} \begin{pmatrix} \cos(\beta t) & -\sin(\beta t) \\ \sin(\beta t) & \cos(\beta t) \end{pmatrix}.$$

We could in fact make use of the complex eigenvectors to derive the following analogy to the expression for the flow map in the real distinct eigenvalue case:

$$\exp(At) = e^{(\alpha+i\beta)t} P_{\mathbf{v}} + e^{(\alpha-i\beta)t} P_{\overline{\mathbf{v}}},$$

where (with "perpendicular" defined with respect to standard inner product on $\mathbb{C}^2$: $\mathbf{v} \cdot \mathbf{w} = \sum_i v_i \overline{w_i}$ with $\mathbf{v} = \sum_i v_i e_i$, $\mathbf{v} = \sum_i w_i e_i$ and $\{e_i\}$ is an orthonormal basis)

$$P_{\mathbf{v}} = \mathbf{v}(\overline{(\overline{\mathbf{v}}^{\perp})}^{\top} \mathbf{v})^{-1} \overline{(\overline{\mathbf{v}}^{\perp})}^{\top}, \quad P_{\overline{\mathbf{v}}} = \overline{P_{\mathbf{v}}}$$

Example

In the case of (8) the eigenvector of A for eigenvalue $\lambda = \alpha + i\beta$ is $\mathbf{v} = \begin{pmatrix} 1 \\ -i \end{pmatrix}$, $\overline{\mathbf{v}} = \begin{pmatrix} 1 \\ i \end{pmatrix}$, and we choose $\overline{\mathbf{v}}^{\perp} = \begin{pmatrix} i \\ 1 \end{pmatrix}$. Then we obtain

$$P_{\mathbf{v}} = \mathbf{v}(\overline{(\overline{\mathbf{v}}^{\perp})}^{\top}\mathbf{v})^{-1}\overline{(\overline{\mathbf{v}}^{\perp})}^{\top} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2i} \\ \frac{1}{2i} & \frac{1}{2} \end{pmatrix} \text{ so that}$$

$$\begin{aligned} \exp((\alpha+i\beta)t)P_{\mathbf{v}} + c.c. &= e^{\alpha t}\left[\begin{pmatrix} \frac{1}{2}e^{i\beta t} & -\frac{1}{2i}e^{i\beta t} \\ \frac{1}{2i}e^{i\beta t} & \frac{1}{2}e^{i\beta t} \end{pmatrix} + c.c.\right] \\ &= e^{\alpha t}\begin{pmatrix} \cos(\beta t) & -\sin(\beta t) \\ \sin(\beta t) & \cos(\beta t) \end{pmatrix}. \end{aligned}$$

Example

Let $A = \begin{pmatrix} 1 & -3 \\ 2 & 1 \end{pmatrix}$, then one eigenvalue is equal to $1 + i\sqrt{6}$ with corresponding eigenvector $\mathbf{v} = \begin{pmatrix} 1 \\ -\frac{i\sqrt{6}}{3} \end{pmatrix}$.

Choosing $\overline{\mathbf{v}}^\perp = \begin{pmatrix} \frac{i\sqrt{6}}{3} \\ 1 \end{pmatrix}$ we obtain

$$P_{\mathbf{v}} = \mathbf{v}((\overline{\overline{\mathbf{v}}^\perp})^\top \mathbf{v})^{-1}(\overline{\overline{\mathbf{v}}^\perp})^\top = \begin{pmatrix} \frac{1}{2} & i\frac{3}{2\sqrt{6}} \\ -i\frac{1}{\sqrt{6}} & \frac{1}{2} \end{pmatrix} \text{ so that}$$

$$\begin{aligned} \exp((1+i\sqrt{6})t)P_{\mathbf{v}} + c.c. &= e^t\left[e^{i\sqrt{6}t}\begin{pmatrix} \frac{1}{2} & -\frac{1}{2i}\frac{\sqrt{6}}{2} \\ \frac{1}{2i}\frac{\sqrt{6}}{3} & \frac{1}{2} \end{pmatrix} + c.c.\right] \\ &= e^t\begin{pmatrix} \cos(t\sqrt{6}) & -\frac{\sqrt{6}}{2}\sin(t\sqrt{6}) \\ \frac{\sqrt{6}}{3}\sin(t\sqrt{6}) & \cos(t\sqrt{6}) \end{pmatrix}. \end{aligned}$$

Some preliminaries from linear algebra:

Definition (Algebraic multiplicity)

If a polynomial can be written as $p(r) = (r - \lambda)^k q(r)$ with $q(\lambda) \neq 0$ then λ is called a root of p of algebraic multiplicity k.

In the calculation of eigenvalues, the terminology "algebraic multiplicity of an eigenvalue" is used with reference to the characteristic polynomial obtained from $\det(A - \lambda I) = 0$.

Definition (Geometric multiplicity of eigenvalues)

An eigenvalue has geometric multiplicity k if the eigenspace associated to this eigenvalue has dimension k.

In case the algebraic and geometric multiplicities of an eigenvalue are not equal to each other, it is useful to introduce the notion of a generalized eigenspace.

Definition (Generalized eigenspace)

The generalized eigenspace E_λ for eigenvalue λ of a square matrix A, with algebraic multiplicity n, is equal to $\ker((A - \lambda I)^n)$.

The following observation generalizes the fact that eigenspaces are flow invariant.

Proposition

Generalized eigenspaces of $A \in gl(m, \mathbb{R})$ are flow invariant for the ODE $\frac{d}{dt}x = Ax$, $x \in \mathbb{R}^m$.

Proof.

Let n denote the algebraic multiplicity of an eigenvalue λ of A. Let $\mathbf{w} \in E_\lambda$ (the generalized eigenspace). Then, since $(A - \lambda I)^n \mathbf{w} = 0$, $A\mathbf{w}$ satisfies also the property that $A\mathbf{w} \in \ker(A - \lambda I)^n$. This shows that solutions with initial conditions in E_λ remain within E_λ. □

Let us consider now the general case of a matrix $A \in gl(2, \mathbb{R})$ with a real eigenvalue λ of algebraic multiplicity two and geometric multiplicity one.[1] Let $\mathbf{v}$ denote the eigenvector. The generalized eigenspace E_λ is spanned by $\mathbf{v}$ and another vector, which we call $\mathbf{w}$, where $\mathbf{w} \notin \ker(A - \lambda I)$. Since $\mathbf{w} \in \ker(A - \lambda I)^2$ it follows that $(A - \lambda I)\mathbf{w} \in \ker(A - \lambda I)$ so that

$$A\mathbf{w} = \lambda \mathbf{w} + c\mathbf{v},$$

for some $c \neq 0$ (otherwise $\mathbf{w} \in \ker(A - \lambda I)$). By scaling of $\mathbf{w}$ one can set $c = 1$ without loss of generality. Writing vectors in the basis $\{\mathbf{v}, \mathbf{w}\}$, $x(t) = v(t)\mathbf{v} + w(t)\mathbf{w}$, we are left to solve

$$\frac{d}{dt}v(t) = \lambda v(t) + w(t), \quad \frac{d}{dt}w(t) = \lambda w(t)$$

yielding the solution, for initial condition $(v(0), w(0)) = (\mathbf{v}_0, \mathbf{w}_0)$,

$$\mathbf{v}(t) = e^{\lambda t}\mathbf{v}_0 + te^{\lambda t}\mathbf{w}_0, \quad \mathbf{w}(t) = e^{\lambda t}\mathbf{w}_0.$$

With this basis, $A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ and $\Phi^t = \begin{pmatrix} e^{\lambda t} & te^{\lambda t} \\ 0 & e^{\lambda t} \end{pmatrix}$.

From the above described methodology it already transpired that the problem decomposes to ODEs on generalized eigenspaces. On each of these eigenspaces one can choose convenient coordinates. We have already seen some examples:

- If $\lambda \in \mathbb{R}$ and the geometric multiplicity of λ is equal to the algebraic multiplicity, the eigenvectors for λ span E_λ. Choosing these as a basis yields a diagonal matrix A.
- If $\lambda \notin \mathbb{R}$ and the geometric multiplicity of λ is equal to the algebraic multiplicity, the linear combinations of (complex) eigenvectors $(\mathbf{v} + \overline{\mathbf{v}})$, $i(\mathbf{v} - \overline{\mathbf{v}})$ span E_λ. Choosing these as a basis yields a blockdiagonal matrix A with 2×2 blocks of the form (8).

In case the geometric multiplicity of eigenvalues is less than the algebraic multiplicity, we have so-called *Jordan blocks* (hence the use of this terminology in the above example). You should be familiar with these (from M2P2) in the case of matrices with complex coefficients. For real matrices, analogous results can be proven, and Jordan blocks take the form

$$\lambda \text{ real} \quad : \quad \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix}$$

$$\lambda = \alpha \pm i\beta \text{ complex} \quad : \quad \begin{pmatrix} R_\lambda & I & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & R_\lambda & I \\ 0 & \cdots & \cdots & 0 & R_\lambda \end{pmatrix}, \quad \text{and} \quad R_\lambda = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix},$$

and I denotes the 2×2 identity matrix.

The Jordan normal form arises naturally as a consequence of a convenient choice of basis for the coordinates. Alternatively, we can also take the viewpoint that any matrix A can be transformed into Jordan normal form by an appropriate linear coordinate transformation.
Suppose $e_1, \ldots, e_m$ is a basis for $\mathbb{R}^m$ in terms of which we have written our ODE. We consider a new basis $\hat{e}_1, \ldots, \hat{e}_m$ obtained by a linear transformation $T \in Gl(m, \mathbb{R})$ (group of invertible linear maps from $\mathbb{R}^m$ to itself): $\hat{e}_j = Te_j$. We consider the consequence for the matrix representation of the linear vector field. We can write points in $\mathbb{R}^m$ in two ways, as $\sum_{i=1}^m x_i e_i = \sum_{i=1}^m y_i \hat{e}_i$. Let $x = (x_1, \ldots, x_m)$ and $y = (y_1, \ldots, y_m)$. Then x and y are related as $x = Ty$. For the ODE this means

$$\frac{dx}{dt} = Ax \Leftrightarrow \frac{d(Ty)}{dt} = ATy \Leftrightarrow \frac{dy}{dt} = \tilde{A}y, \quad \text{where} \quad \tilde{A} = T^{-1}AT.$$

So the Jordan normal form result implies that there exists a linear coordinate transformation represented by $T \in Gl(m, \mathbb{R})$ that conjugates A to $\tilde{A}$ where the latter has the structure of the (real) Jordan normal form. Often, if there is no specific significance attached to any particular choice of coordinates, it is assumed that linear vector fields are in Jordan normal form.

Definition (Semi-simple and nilpotent linear maps)

A linear map $A \in gl(m, \mathbb{R})$ is called *semi-simple* if the geometric and algebraic multiplicities of each eigenvalue are equal to each other. A linear map $A \in gl(m, \mathbb{R})$ is called *nilpotent* if there exists a $p \in \mathbb{N}$ such that $A^p = 0$ (the null matrix).

Definition (Jordan-Chevalley decomposition)

Let $A \in gl(m, \mathbb{R})$. Then $A = S + N$ is called the *Jordan-Chevalley decomposition* of A if S is semi-simple, N is nilpotent and N and S commute, i.e. $NS = SN$.

Theorem

Every $A \in gl(m, \mathbb{R})$ admits a unique Jordan-Chevalley decomposition.

Please note that the above notions and results also hold for complex matrices in $gl(m, \mathbb{C})$.

Note that the Jordan-Chevalley decomposition of a matrix in Jordan normal form consists of the "diagonal" (semi-simple) and "off-diagonal" (nilpotent) parts (note that indeed these parts are semi-simple and nilpotent, respectively, and that they commute).
If $A = N + S$ is the Jordan-Chevalley decomposition, then

$$\exp(At) = \exp(St)\exp(Nt),$$

so that one can focus on the flows of the linear ODEs with semi-simple and nilpotent matrices separately. In this respect we note that the exponential $\exp(Nt)$ where N is nilpotent, with p least such that $N^p = 0$, takes the form

$$\exp(Nt) = \sum_{k=0}^{\infty} \frac{(Nt)^k}{k!} = \sum_{k=0}^{p-1} \frac{(Nt)^k}{k!}$$

and thus is polynomial in t of degree $p - 1$.

Example

The Jordan-Chevalley decomposition of the matrix $A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ is

$$A = S + N, \text{ where } S = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \text{ and } N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Consequently, we have

$$\exp(St) = \begin{pmatrix} e^{\lambda t} & 0 \\ 0 & e^{\lambda t} \end{pmatrix} \text{ and } \exp(Nt) = I + Nt = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$$

so that

$$\exp(At) = \begin{pmatrix} e^{\lambda t} & 0 \\ 0 & e^{\lambda t} \end{pmatrix} \cdot \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} e^{\lambda t} & te^{\lambda t} \\ 0 & e^{\lambda t} \end{pmatrix}.$$

Example (from 2008 exam)

Let

$$A = \begin{pmatrix} 6 & 1 \\ -4 & 10 \end{pmatrix}.$$

Determine the *real Jordan form* and *Jordan-Chevalley decomposition* of A, and the flow of $\dot{x} = Ax$ with $x \in \mathbb{R}^2$.

- eigenvalues: $(6-\lambda)(10-\lambda)+4 = (8-\lambda)^2$. $\lambda = 8$ with algebraic multiplicity 2.
- only one-dimensional eigenspace: $\langle \begin{pmatrix} 1 \\ 2 \end{pmatrix} \rangle$, so geometric multiplicity 1
- Jordan normal form $J = \begin{pmatrix} 8 & 1 \\ 0 & 8 \end{pmatrix}$
- Jordan-Chevalley decomposition of Jordan normal form, $J = S + N$ with $S = \begin{pmatrix} 8 & 0 \\ 0 & 8 \end{pmatrix}$ and $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

Example (cont)

- Jordan-Chevaley decomposition of A: note that $J = P^{-1}AP$ for some $P \in Gl(2,\mathbb{R})$, so semi-simple part of A_s of A is PSP^{-1} and nilpotent part A_n of A is PNP^{-1}. But since S is a multiple of I, we have $PSP^{-1} = S$, so $A_s = S$ and $A_n = A - A_s = \begin{pmatrix} -2 & 1 \\ -4 & 2 \end{pmatrix}$. (verify that $A_n^2 = 0$)
- $\exp(At) = e^{8t}(\mathrm{I} + A_n t) = e^{8t}\begin{pmatrix} 1-2t & t \\ -4t & 1+2t \end{pmatrix}$

Linear autonomous ODEs always have an equilibrium at 0, since there we have $\frac{dx}{dt} = A0 = 0$, and thus 0 is always a fixed point of the flow $\Phi^t = \exp(At)$. One of the elementary things one may want to understand about dynamics is whether points in the neighbourhood of an equilibrium have a tendency to stay nearby, or whether they have a tendency to wander off. We call the former type *stable*. The formal definition is:

Definition (Lyapunov and asymptotic stability)

An equilibrium or fixed point $\bar{x}$ is Lyapunov stable, if for every neighbourhood U of $\bar{x}$, there exists neighbourhoods V_0 and V_1 of $\bar{x}$ such that $V_1 \subset V_0 \subset U$ and the forward time orbits of all initial conditions $x_0 \in V_1$ remain in V_0 for all time.
A special case of Lyapunov stability is *asymptotic stability*, which require that all initial conditions in V_1 converge to $\bar{x}$ as time goes to infinity.

Certain properties of the eigenvalues of a linear ODE are directly related to stabilty.

Proposition

Let $A \in gl(m, \mathbb{R})$. Then the trivial fixed point 0 of the flow $\Phi^t = \exp(At)$ is not Lyapunov stable if the real part of one eigenvalue of A is positive. The fixed point is asymptotically stable if and only if the real parts of all eigenvalues of A are negative.

Proof.

We first note that (generalized) eigenspaces are flow invariant. Whenever there is an eigenvalue with positive real part, apart from the trivial fixed point, all other initial conditions within this eigenspace tend to infinity as time goes forward, which contradicts stability. If all eigenvalues have a negative real part, the flow on each (generalized) eigenspace is contracting (towards 0) which implies asymptotic stability. If there exists an eigenvalue with zero linear part, then the flow on the corresponding eigenspace preserves the distance of initial conditions to 0 which contradicts asymptotic stability. □

Consider a linear autonomous ODE $\frac{dx}{dt} = Ax$ $A \in gl(m, \mathbb{R})$ and $x \in \mathbb{R}^m$, then we may write the phase space $\mathbb{R}^m$ as a direct sum of generalized eigenspaces E_λ (where we take the convention that in case $\lambda \notin \mathbb{R}$, E_λ denotes the invariant subspace associated with eigenvalues λ and $\overline{\lambda}$):

$$\mathbb{R}^m = \oplus_j E_{\lambda_j}.$$

(recall that a sum of vector spaces is a direct sum if all of the components in the sum only have trivial intersection).
We define the *stable* and *unstable* subspaces of $\mathbb{R}^m$ as the union of generalized eigenspaces for eigenvalues with negative and positive real parts, respectively:

$$E^s = \oplus_{\mathrm{Re}(\lambda_j)<0} E_{\lambda_j}, \quad E^u = \oplus_{\mathrm{Re}(\lambda_j)>0} E_{\lambda_j}.$$

The union of the remaining generalized eigenspaces (for eigenvalues with zero real part) is called the *centre subspace*:

$$E^c = \oplus_{\mathrm{Re}(\lambda_j)=0} E_{\lambda_j}.$$

The flow on the stable subspace is asymptotically stable, and the flow on the unstable subspace is asymptotically unstable (the inverse flow is asymptotically stable). The flow on the center manifold is not asymptotically stable (or unstable) as there is always a subspace on which the flow is either stationary (corresponding to zero eigenvalue) or a rotation (purely imaginary complex conjugate pair of eigenvalues). The centre subspace E^c is Lyapunov stable in case there are no nontrivial Jordan blocks.

In general it is not easy to prove that an equilibrium is Lyapunov Stable or asymptotically stable, but we will see a few instances where it is possible. In the exercises you find some discussion about the Lyapunov functions, which can be used to prove stability. The difficulty with that method is that there is no general way of constructing those Lyapunov functions, so that one in the end here needs to rely on some particular insight or intuition in order to construct a Lyapunov function.

Finally we give some simple examples of Lyapunov stable equilibria.

Example

Consider $\frac{dx}{dt} = Ax$ with $x \in \mathbb{R}^2$ and

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

Then $x = 0$ is an equilibrium point. The eigenvalues of A are $a \pm i$. If $a \leq 0$ the equilibrium is Lyapunov stable. We readily compute that $|\Phi^t x| = e^{at}|x|$ for all $x \in \mathbb{R}^2$. In case $a = 0$ then the flow is a rotation around the origin and whereas the equilibrium is Lyapunov stable, it is not asymptotically stable. If $a < 0$ the equilibrium is asymptotically stable.

We finally consider some geometrical properties of the flow that apply to linear as well as nonlinear autonomous ODEs.
We recall that the right hind side of the ODE

$$\frac{dx}{dt} = f(x), \quad x \in \mathbb{R}^m \tag{9}$$

provides for each point in the phase space the *vector* $f(x)$; accordingly we call f a *vector field*. The ODE further specifies that we identify the vector $f(x)$ with dx/dt. The following property is essential:

Proposition

Let $x(t)$ be a solution of (9). Then for each t_0, the vector $f(x(t_0))$ is tangent to $x(t)$ at $x(t_0)$.

We will present the proof for a planar ODE ($m = 2$). The arguments are similar when $m > 2$.
First we find an expression for the tangent vector. Suppose that the curve $x(t)$ has the implicit form $F(x, y) = 0$. To find the tangent in $(x_0, y_0) \in x(t)$ we evaluate the function F in the point $(x_0 + \delta x, y_0 + \delta y)$ with δx and δy small, and write the Taylor series expansion

$$\begin{aligned} F(x_0 + \delta x, y_0 + \delta y) &= F(x_0, y_0) + \frac{\partial F}{\partial x}(x_0, y_0)\delta x + \\ &\quad \frac{\partial F}{\partial y}(x_0, y_0)\delta y + O(|\delta x, \delta y|^2). \end{aligned}$$

Thus in order to stay on the curve, this expression needs to be equal to zero. We already have $F(x_0, y_0) = 0$ since (x_0, y_0) is on the curve, so to first order we need to satisfy

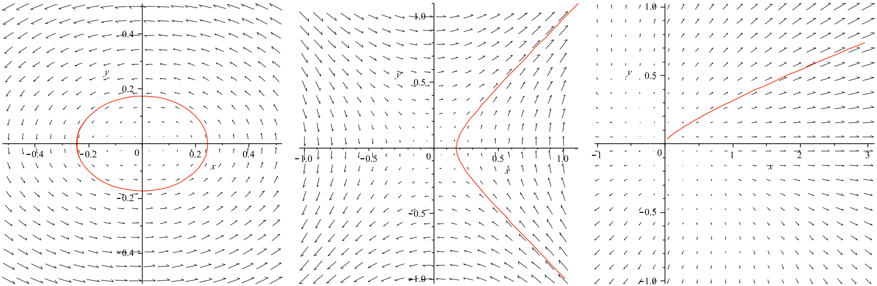
$$\begin{pmatrix} \frac{\partial F}{\partial x}(x_0, y_0) \\ \frac{\partial F}{\partial y}(x_0, y_0) \end{pmatrix} \cdot \begin{pmatrix} \delta x \\ \delta y \end{pmatrix} = 0$$

The vector on the left hand side in this inner product is known as the *gradient* $\nabla F(x_0, y_0)$. The vector on the right hand side is the best linear approximation to the curve in (x_0, y_0) which is known as the *tangent*. We thus observe that the tangent vector is perpendicular to the gradient of the function that defines the curve.

We finally relate this to the ODE. Taking the derivative of F with respect to t, we obtain

$$\frac{d}{dt}F(x_0, y_0) = \frac{\partial F}{\partial x}(x_0, y_0)\dot{x} + \frac{\partial F}{\partial y}(x_0, y_0)\dot{y} = \nabla F(x_0, y_0) \cdot \dot{\mathbf{x}}.$$

Thus $\dot{\mathbf{x}} = f(\mathbf{x})$ is also tangent to the solution curve $\mathbf{x}(t)$.



Examples of linear planar vector fields [made with Maple] in the (from left to right) elliptic, hyperbolic and Jordan block cases. Part of one solution curve is also plotted to illustrate the fact that such curves are everywhere tangent to the vector field. In the case of planar vectors fields, we often provide a sketch the flow by drawing some representative set of solution curves, with arrows on them indicating the direction of the flow. Such a sketch is called a phase portrait.