Differential Equations 2020/21
MA 209

Extra Notes 1

Calculating the exponential of a matrix
Eigenvalues and eigenvectors
Classification of 2×2 matrices
Decoupled systems
Higher dimensional linear systems

1.1 Calculating the exponential of a matrix

As described in the Lecture Notes, if A is $n \times n$ real matrix, then we define the exponential e^A by

$$e^A = I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots = \sum_{k=0}^{\infty} \frac{1}{k!}A^k.$$

But actually calculating this exponential is usually far from trivial. In these extra notes we have a closer look at one method that might be useful for this.

- Suppose we can write $A = PMP^{-1}$, for some invertible matrix P and any other matrix M. Then we have that

$$\begin{aligned} A^k &= (PMP^{-1})^k = (PMP^{-1})\,(PMP^{-1}) \cdots (PMP^{-1}) \\ &= PMP^{-1}\,PMP^{-1} \cdots PMP^{-1} = P\,M\,M \cdots M\,P^{-1} = PM^kP^{-1}, \end{aligned}$$

and hence

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!}A^k = \sum_{k=0}^{\infty} \frac{1}{k!}PM^kP^{-1} = P\,\Big(\sum_{k=0}^{\infty} \frac{1}{k!}M^k\Big)\,P^{-1} = Pe^MP^{-1}.$$

(In fact, we also find that $e^{tA} = Pe^{tM}P^{-1}$; check for yourself.)

So if we can find a matrix M so that $A = PMP^{-1}$ for some invertible matrix P, and e^M is "easy" to determine, then we also can determine e^A.

- One easy type of matrices are diagonal matrices: $D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}$. Then we find for the powers $D^k = \begin{bmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n^k \end{bmatrix}$, and hence

$$e^D = \sum_{k=0}^{\infty} \frac{1}{k!} D^k = \sum_{k=0}^{\infty} \frac{1}{k!} \begin{bmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n^k \end{bmatrix}$$

$$= \begin{bmatrix} \sum\limits_{k=0}^{\infty} \frac{1}{k!}\lambda_1^k & 0 & \cdots & 0 \\ 0 & \sum\limits_{k=0}^{\infty} \frac{1}{k!}\lambda_2^k & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sum\limits_{k=0}^{\infty} \frac{1}{k!}\lambda_n^k \end{bmatrix} = \begin{bmatrix} e^{\lambda_1} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & e^{\lambda_n} \end{bmatrix}.$$

- In the remainder of these extra notes we will show that every real matrix A has some standard accompanying real matrix J so that $A = PJP^{-1}$ for some real invertible matrix P and so that e^J is not too hard to determine. This special matrix is called the *Jordan form* of A.

  We prove that each 2×2 matrix has a Jordan form in quite some detail. For larger dimensions we give the outcomes only and some ideas where it comes from.

  Before we can start looking at the Jordan form, we have a closer look at eigenvalues and eigenvectors of real matrices.

1.2 Eigenvalues

A real or complex number λ is an *eigenvalue* of an $n \times n$ square matrix A if there exists a non-zero vector v so that $Av = \lambda v$.

Since this definition is equivalent to asking for $(A - \lambda I)v = 0$, which only has a non-zero solution if $A - \lambda I$ is singular (i.e. has no inverse), we find all eigenvalues by looking for solutions of the equation $\det(A - \lambda I) = 0$.

- Here is something you should know:

  Theorem 1.1
  An $n \times n$ matrix A has n eigenvalues $\lambda_1, \ldots, \lambda_n$, which can be complex numbers, and where the same number can appear more than once.

  "Proof" As mentioned above, the eigenvalues of A are just the solutions to $\det(A - \lambda I) = 0$. If we define the function $f(\lambda) = \det(A - \lambda I)$ for $\lambda \in \mathbb{C}$, then it's straightforward to show that $f(\lambda)$ is a polynomial of degree n. So the eigenvalues of A are exactly the roots of this degree n polynomial $f(\lambda)$.

  So we need that a polynomial of degree n has n roots. This is fairly hard to prove completely, so we just rely on the Fundamental Theorem of Algebra (FToA):

* **Fundamental Theorem of Algebra**
Let $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ be a polynomial over the complex numbers of degree $n \geq 1$ (hence $a_n \neq 0$). Then $p(x)$ has n roots $r_1, \ldots, r_n$ in the complex numbers (and can be written as $p(x) = a_n(x - r_1)(x - r_2) \cdots (x - r_n)$).

Translating this back to our polynomial $f(\lambda) = \det(A - \lambda I)$, we get that $f(\lambda)$ has exactly n roots, hence A has exactly n eigenvalues. ■

- Note that the eigenvalues can be complex numbers. Also note that the same eigenvalue can appear more than once in the list $\lambda_1, \ldots, \lambda_n$. The number of times the same number appears is called its *multiplicity*.

  We sometimes emphasise this by listing only the different eigenvalues $\lambda_1, \ldots, \lambda_k$ (where we must have $k \leq n$) with their multiplicities $m_1, \ldots, m_k$. These multiplicities m_i are positive integers with $m_1 + \cdots + m_k = n$.

- Theorem 1.1 doesn't say much about what kind of numbers we can expect the eigenvalues to be. And in general, there isn't much that can be said. But in the case we are interested in in this course we can do a little more:

  Theorem 1.2
  Suppose A is an $n \times n$ matrix in which all entries are real numbers.

  (a) If A has a complex eigenvalue $\lambda = \alpha + \beta i$, with $\alpha, \beta \in \mathbb{R}$, $\beta \neq 0$, then the conjugate of λ, $\overline{\lambda} = \alpha - \beta i$, is also an eigenvalue of A.

  (b) If n is odd, then at least one of the eigenvalues of A is a real number.

  "Proof" We again look at the polynomial $f(\lambda) = \det(A - \lambda I)$ of degree n, realising that the eigenvalues are exactly the roots of this polynomial. Suppose $f(\lambda) = a_n\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$. If all entries of A are real, then so are those of $A - \lambda I$, and hence when writing out the determinant $\det(A - \lambda I)$ we also only see real numbers. So we know that all coefficients $a_n, \ldots, a_0$ of the polynomial $f(\lambda)$ are real numbers.

  Now it's a matter of writing out in full the expressions for both $f(\alpha + \beta i)$ and $f(\alpha - \beta i)$, and in particular looking at their real and imaginary part (using that the coefficients $a_n, \ldots, a_0$ of the polynomial $f(\lambda)$ are reals). If you do this correctly, you will find that $\mathrm{Re}(f(\alpha+\beta i)) = \mathrm{Re}(f(\alpha-\beta i))$, while $\mathrm{Im}(f(\alpha + \beta i)) = -\mathrm{Im}(f(\alpha - \beta i))$. Since $\alpha + \beta i$ is assumed to be an eigenvalue of A, we have $f(\alpha + \beta i) = 0$, hence $\mathrm{Re}(f(\alpha + \beta i)) = 0$ and $\mathrm{Im}(f(\alpha + \beta i)) = 0$. This means that also $\mathrm{Re}(f(\alpha - \beta i)) = 0$ and $\mathrm{Im}(f(\alpha - \beta i)) = 0$, and hence $f(\alpha - \beta i) = 0$.

  It follows that $\alpha - \beta i$ is also an eigenvalue of A. This proves part (a).

  We prove part (b) by induction on the degree n of the polynomial $f(\lambda) = \det(A - \lambda I)$. If $n = 1$, then we have $f(\lambda) = a_1\lambda + a_0$, where a_1, a_0 are real numbers and $a_1 \neq 0$. This polynomial has the obvious real root $r = -a_0/a_1$.

  So now suppose $f(\lambda)$ has odd degree $n \geq 3$. If it has no complex root, then it has n real roots, and we are done. So suppose $f(\lambda)$ has a complex root $z = \alpha + \beta i$. Above we've seen that then also the conjugate $\overline{z} = \alpha - \beta i$ is a root. That means we can factor out the factors $\lambda - z$ and $\lambda - \overline{z}$ and write $f(\lambda) = (\lambda - z)(\lambda - \overline{z})\, q(\lambda)$, where $q(\lambda)$ is a polynomial of degree $n - 2$. If we multiply out we get (use that $i^2 = -1$)

$$\begin{aligned}(\lambda - z)(\lambda - \overline{z}) &= (\lambda - \alpha - \beta i)(\lambda - \alpha + \beta i) \\ &= \lambda^2 - \alpha\lambda + \beta\lambda i - \alpha\lambda + \alpha^2 - \alpha\beta i - \beta\lambda i + \alpha\beta i - \beta^2 i^2 \\ &= \lambda^2 - 2\alpha\lambda + \alpha^2 + \beta^2.\end{aligned}$$

So in fact we have $f(\lambda) = (\lambda^2 - 2\alpha\lambda + \alpha^2 + \beta^2)\, q(\lambda)$. Since all coefficients of $f(\lambda)$ and all of $\alpha, \alpha^2, \beta^2$ are real numbers, also all coefficients of $q(\lambda)$ are real numbers. And since $q(\lambda)$ has odd degree $n-2$, by induction we know it has a real root. But any root of $q(\lambda)$ is also a root of $f(\lambda)$. ■

1.3 Eigenvectors

If λ is an eigenvalue of an $n \times n$ square matrix A, then an *eigenvector* is a non-zero vector v so that $Av = \lambda v$.

Theorem 1.3
Let λ be an eigenvalue of multiplicity m. Then the number of linearly independent eigenvectors of λ is between 1 and m.

"Proof" We only show that there is at least one eigenvector of λ. The proof that there can't be more than the multiplicity would involve things that go beyond what we want to know in this course. Recall the fact that λ is an eigenvalue means that $\det(A - \lambda I) = 0$. Let $c_1, c_2, \ldots, c_n$ be the columns of the matrix $A - \lambda I$. The fact that the determinant of $A - \lambda I$ is zero means that the columns form a dependent set. Hence there exist numbers $a_1, \ldots, a_n$, not all equal to 0, so that $a_1c_1 + a_2c_2 + \cdots + a_nc_n = 0$. Now let v be the vector that has the numbers $a_1, \ldots, a_n$ as entries: $v = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$. Since at least one from $a_1, \ldots, a_n$ is not 0, we have $v \neq 0$. Writing writing out the product, we get $(A - \lambda I)v = a_1c_1 + a_2c_2 + \cdots + a_nc_n = 0$. This is the same as $Av - \lambda v = 0$. Hence $Av = \lambda v$, and we found an eigenvector for λ. ■

Theorem 1.4
Let v_1 be an eigenvector of an eigenvalue λ_1 and v_2 an eigenvector of an eigenvalue λ_2, where $\lambda_1 \neq \lambda_2$. Then v_1, v_2 are linearly independent vectors.

Proof Exercise. ■

- Since we allow eigenvalues to be complex numbers, it seems also possible that entries of eigenvectors can be complex numbers. If this is the case, then we often will split the entries in their real and imaginary part. Hence we will write $v = v_R + iv_I$, where all entries in both v_R and v_I are real numbers.

  As the following result shows, if a real matrix has a complex eigenvalue, then we always have a complex eigenvector with additional properties.

Theorem 1.5
Let A be an $n \times n$ matrix in which all entries are real numbers. Suppose A has a complex eigenvalue $\lambda = \alpha + \beta i$, with $\alpha, \beta \in \mathbb{R}$, $\beta \neq 0$, with eigenvector v.

(a) The eigenvector v contains complex entries. I.e. if we write $v = v_R + iv_I$, then $v_I \neq 0$.

(b) In fact, if we write $v = v_R + iv_I$, then the two parts v_R, v_I are two linearly independent vectors.

(c) The parts v_R, v_I satisfy $Av_R = \alpha v_R - \beta v_I$ and $Av_I = \beta v_R + \alpha v_I$.

Proof For (a), suppose there is an eigenvector v of $\lambda = \alpha + \beta i$ in which all entries are real numbers. Then in the expression Av we have real numbers only. On the other hand, the expression $\lambda v = (\alpha + \beta i)v$ has numbers with imaginary parts as well. But since we must have $Av = \lambda v$, this can't be correct. Hence v cannot contain real entries only.

Part (b) is again an exercise.

For part (c), first use the linearity of matrix multiplication to write $Av = A(v_R+iv_I) = Av_R+iAv_I$. Here everything in Av_R and Av_I are real numbers. Next write $\lambda v = (\alpha+\beta i)(v_R+iv_I) = (\alpha v_R - \beta v_I) + i(\alpha v_I + \beta v_R)$. Since we have $Av = \lambda v$ and both real and imaginary parts must be equal, we have $Av_R = \alpha v_R - \beta v_I$ and $Av_I = \alpha v_I + \beta v_R$. So we are done. ∎

1.4 Classification of 2-dimensional matrices

With the knowledge from above we are now ready to describe all different possibilities for the eigenvalues and eigenvectors of a real-valued 2×2 matrix A:

(i) A has two different real eigenvalues λ_1, λ_2, $\lambda_1 > \lambda_2$, each with an eigenvector v_1, v_2 (Since the two λs are different, we can always number them so that $\lambda_1 > \lambda_2$.);

(ii) A has a double real eigenvalue λ, with two linearly independent eigenvectors v_1, v_2;

(iii) A has a double real eigenvalue λ, with only one eigenvector v;

(iv) A has a complex eigenvalue $\lambda = \alpha+\beta i$, $\beta > 0$, with an eigenvector $v = v_R + iv_I$.
(Note that by the Theorem 1.2 we know that if $\lambda = \alpha+\beta i$ is a complex eigenvalue, then so is $\overline{\lambda} = \alpha - \beta i$. Since one of $\beta, -\beta$ is positive, we loose nothing by assuming $\beta > 0$.)

We will consider each of these four cases separately below.

- (i) Let P be the 2×2 matrix formed by using the eigenvectors v_1, v_2 as columns; so we can write $P = \big[v_1 \,|\, v_2\big]$. This means that $AP = A\big[v_1 \,|\, v_2\big] = \big[\,\lambda_1 v_1 \,|\, \lambda_2 v_2\,\big]$. This answer can also be written as $\big[v_1 \,|\, v_2\big]\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}$, which is the same as $P\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}$. So we find that $AP = P\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}$. Since v_1, v_2 are linearly independent the matrix $P = \big[v_1 \,|\, v_2\big]$ is invertible. Multiplying from the right by P^{-1} we get the following:

  * *The invertible real matrix* $P = \big[v_1 \,|\, v_2\big]$ *has the property that* $A = P\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}P^{-1}$.

- (ii) In the case that we have a double eigenvalue λ with two linearly independent eigenvectors v_1, v_2 we can do exactly as above. The conclusion will be:

  * *The invertible real matrix* $P = \big[v_1 \,|\, v_2\big]$ *has the property that* $A = P\begin{bmatrix}\lambda & 0\\ 0 & \lambda\end{bmatrix}P^{-1}$.

- (iii) This case is in a sense the most complicated case, since we only have one vector yet. To get a second one, we would need to go deeper into the theory of eigenvalues; deeper than we like to do at the moment. The eigenvector v of A for eigenvalue λ is the only solution for x (up to scalar multiplication) of the equation $(A-\lambda I)x = 0$. Since λ is a double eigenvalue, it can be shown that the equation $(A-\lambda I)^2 x = 0$ has two linearly independent solutions. Since $(A-\lambda I)^2 v = (A-\lambda I)(A-\lambda I)v = (A-\lambda I)0 = 0$, we can take v as one of these solutions. Suppose w^* is a second one, where v, w^* are linearly independent. Since $(A-\lambda I)^2 w^* = (A-\lambda I)(A-\lambda I)w^* = 0$, it follows that the vector $u = (A-\lambda I)w^*$ has the property that $(A-\lambda I)u = 0$. But the only linearly independent vector for which $(A-\lambda I)x = 0$ was v, so u must be a multiple of v. Say we have $u = kv$. Since $k \neq 0$ (otherwise $u = 0$ and $(A-\lambda I)w^* = u = 0$ would be another solution to $(A-\lambda I)x = 0$), we can write $w = \frac{1}{k}w^*$. So we have that $(A-\lambda I)w = (A-\lambda I)\frac{1}{k}w^* = \frac{1}{k}(A-\lambda I)w^* = \frac{1}{k}u = v$. This is the same as $Aw = v + \lambda w$. Since v, w^* are linearly independent and $w = \frac{1}{k}w^*$, also v, w are linearly independent.

  Now let P be the 2×2 matrix formed by using the vectors v, w as columns: $P = \big[v \,|\, w\big]$. This means that $AP = A\big[v \,|\, w\big] = \big[\,\lambda v \,|\, v + \lambda w\,\big]$. This answer can also be written as $\big[v \,|\, w\big]\begin{bmatrix}\lambda & 1\\ 0 & \lambda\end{bmatrix}$, which is the same as $P\begin{bmatrix}\lambda & 1\\ 0 & \lambda\end{bmatrix}$. So we find that $AP = P\begin{bmatrix}\lambda & 1\\ 0 & \lambda\end{bmatrix}$. Since v, w are linearly independent, the matrix $P = \big[v \,|\, w\big]$ is invertible. Multiplying from the right by P^{-1} we get:

* *The invertible real matrix* $P\big[v \,|\, w\big]$ *has the property that* $A = P\begin{bmatrix}\lambda & 1\\ 0 & \lambda\end{bmatrix}P^{-1}$.

- (iv) We actually could do the case of two complex eigenvalues (who must be different, because one of them is the conjugate of the other) just as Case (i). But that would mean that we start looking at matrices with complex entries. And when taking the exponential it's not clear what would happen with those complex numbers. Moreover, once we translate this whole business to solutions of systems of linear equations, we really don't want to end up with complex numbers in our answers. So we treat this case differently from Case (i).

  We use that v_R, v_I are two linearly independent vectors that satisfy $Av_R = \alpha v_R - \beta v_I$ and $Av_I = \beta v_R + \alpha v_I$.

  Let P be the 2×2 matrix formed by using the vectors v_R, v_I as columns: $P = \big[v_R \,|\, v_I\big]$. Then we easily find $AP = A\big[v_R \,|\, v_I\big] = \big[\,\alpha v_R - \beta v_I \,|\, \beta v_R + \alpha v_I\,\big]$. This answer can also be written as $\big[v_R \,|\, v_I\big]\begin{bmatrix}\alpha & \beta\\ -\beta & \alpha\end{bmatrix}$, which is the same as $P\begin{bmatrix}\alpha & \beta\\ -\beta & \alpha\end{bmatrix}$. Since v_R, v_I are two linearly independent vectors, the matrix $P = \big[v_R \,|\, v_I\big]$ is invertible. Multiplying from the right by P^{-1} we get the following:

  * *The invertible real matrix* $P = \big[v_R \,|\, v_I\big]$ *has the property that* $A = P\begin{bmatrix}\alpha & \beta\\ -\beta & \alpha\end{bmatrix}P^{-1}$.

- The special forms $\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}$ (with $\lambda_1 > \lambda_2$), $\begin{bmatrix}\lambda & 0\\ 0 & \lambda\end{bmatrix}$, $\begin{bmatrix}\lambda & 1\\ 0 & \lambda\end{bmatrix}$, and $\begin{bmatrix}\alpha & \beta\\ -\beta & \alpha\end{bmatrix}$ (with $\beta > 0$) from above are called the *Jordan form* of A.

- Two matrices A, B are called *similar* if there is an invertible matrix Q so that $A = Q^{-1}BQ$. It is easy to show (exercise) that this means that if two matrices are similar then they have the same Jordan form.

1.5 Exponentials of Jordan forms

For each of the Jordan forms J in the previous part, it is not so hard to find the exponentials e^J and e^{tJ}.

- (i) If $J = \begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}$, then for $k = 0, 1, 2, \ldots$ we have $J^k = \begin{bmatrix}\lambda_1^k & 0\\ 0 & \lambda_2^k\end{bmatrix}$, and hence

$$e^J = \sum_{k=0}^{\infty}\frac{1}{k!}J^k = \sum_{k=0}^{\infty}\frac{1}{k!}\begin{bmatrix}\lambda_1^k & 0\\ 0 & \lambda_2^k\end{bmatrix} = \begin{bmatrix}\sum\limits_{k=0}^{\infty}\frac{1}{k!}\lambda_1^k & 0\\ 0 & \sum\limits_{k=0}^{\infty}\frac{1}{k!}\lambda_2^k\end{bmatrix} = \begin{bmatrix}e^{\lambda_1} & 0\\ 0 & e^{\lambda_2}\end{bmatrix}.$$

  A similar argument gives that $e^{tJ} = \begin{bmatrix}e^{t\lambda_1} & 0\\ 0 & e^{t\lambda_2}\end{bmatrix}$.

- (ii) If $J = \begin{bmatrix}\lambda & 0\\ 0 & \lambda\end{bmatrix}$, then we get from the previous case that $e^J = \begin{bmatrix}e^{\lambda} & 0\\ 0 & e^{\lambda}\end{bmatrix}$ and $e^{tJ} = \begin{bmatrix}e^{t\lambda} & 0\\ 0 & e^{t\lambda}\end{bmatrix}$.

- (iii) For the Jordan form $J = \begin{bmatrix}\lambda & 1\\ 0 & \lambda\end{bmatrix}$, we immediately calculate e^{tJ}. We have that $tJ = \begin{bmatrix}t\lambda & t\\ 0 & t\lambda\end{bmatrix}$.

  So we obtain $(tJ)^0 = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$; while for the powers $k \geq 1$ we get $(tJ)^k = \begin{bmatrix}(t\lambda)^k & kt(t\lambda)^{k-1}\\ 0 & (t\lambda)^k\end{bmatrix}$

(this can be easily proved using induction on k).

This gives

$$\begin{aligned}
e^{tJ} &= \sum_{k=0}^{\infty} \frac{1}{k!}(tJ)^k = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \sum_{k=1}^{\infty} \frac{1}{k!} \begin{bmatrix} (t\lambda)^k & kt(t\lambda)^{k-1} \\ 0 & (t\lambda)^k \end{bmatrix} \\
&= \begin{bmatrix} 1 + \sum\limits_{k=1}^{\infty} \frac{1}{k!}(t\lambda)^k & t\sum\limits_{k=1}^{\infty} \frac{1}{k!}k(t\lambda)^{k-1} \\ 0 & 1 + \sum\limits_{k=1}^{\infty} \frac{1}{k!}(t\lambda)^k \end{bmatrix} = \begin{bmatrix} \sum\limits_{k=0}^{\infty} \frac{1}{k!}(t\lambda)^k & t\sum\limits_{k=1}^{\infty} \frac{1}{(k-1)!}(t\lambda)^{k-1} \\ 0 & \sum\limits_{k=0}^{\infty} \frac{1}{k!}(t\lambda)^k \end{bmatrix} \\
&= \begin{bmatrix} \sum\limits_{k=0}^{\infty} \frac{1}{k!}(t\lambda)^k & t\sum\limits_{\ell=0}^{\infty} \frac{1}{\ell!}(t\lambda)^{\ell} \\ 0 & \sum\limits_{k=0}^{\infty} \frac{1}{k!}(t\lambda)^k \end{bmatrix} = \begin{bmatrix} e^{t\lambda} & te^{t\lambda} \\ 0 & e^{t\lambda} \end{bmatrix}.
\end{aligned}$$

By taking $t = 1$, this means $e^J = \begin{bmatrix} e^{\lambda} & e^{\lambda} \\ 0 & e^{\lambda} \end{bmatrix}$.

- (iv) Finally, the Jordan form $J = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$. Writing $J_\alpha = \begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix}$ and $J_\beta = \begin{bmatrix} 0 & \beta \\ -\beta & 0 \end{bmatrix}$ we have that $J = J_\alpha + J_\beta$. It is easy to check that J_α and J_β commute ($J_\alpha J_\beta = J_\beta J_\alpha$), and so by Theorem 1.5.2 from the Lecture Notes we have that $e^J = e^{J_\alpha} e^{J_\beta}$.

  For $J_\alpha = \begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix}$ we can argue as in Case (i) or (ii) to get $e^{J_\alpha} = \begin{bmatrix} e^{\alpha} & 0 \\ 0 & e^{\alpha} \end{bmatrix}$.

  For $J_\beta = \begin{bmatrix} 0 & \beta \\ -\beta & 0 \end{bmatrix}$, life is slightly more involved. Using induction on k it is not so hard to find the following expressions for the powers of J_β, for $k = 0, 1, 2, \ldots$.

  If $k = 2m$ is even, then $J_\beta^k = J_\beta^{2m} = \begin{bmatrix} (-1)^m\beta^k & 0 \\ 0 & (-1)^m\beta^k \end{bmatrix} = \begin{bmatrix} (-1)^{k/2}\beta^k & 0 \\ 0 & (-1)^{k/2}\beta^k \end{bmatrix}$,

  if $k = 2m + 1$ is odd, then

  $$J_\beta^k = J_\beta^{2m+1} = \begin{bmatrix} 0 & (-1)^m\beta^k \\ -(-1)^m\beta^k & 0 \end{bmatrix} = \begin{bmatrix} 0 & (-1)^{(k-1)/2}\beta^k \\ -(-1)^{(k-1)/2}\beta^k & 0 \end{bmatrix}.$$

  So we find that $e^{J_\beta} = \begin{bmatrix} p(\beta) & q(\beta) \\ -q(\beta) & p(\beta) \end{bmatrix}$, where

  $$\begin{aligned}
  p(\beta) &= \sum_{k=0,\ k \text{ even}}^{\infty} \frac{1}{k!}(-1)^{k/2}\beta^k = 1 - \frac{1}{2!}\beta^2 + \frac{1}{4!}\beta^4 - \frac{1}{6!}\beta^6 + \cdots = \cos(\beta), \\
  q(\beta) &= \sum_{k=0,\ k \text{ odd}}^{\infty} \frac{1}{k!}(-1)^{(k-1)/2}\beta^k = \beta - \frac{1}{3!}\beta^3 + \frac{1}{5!}\beta^5 - \frac{1}{7!}\beta^7 + \cdots = \sin(\beta).
  \end{aligned}$$

  This leads to

  $$e^J = e^{J_\alpha} \cdot e^{J_\beta} = \begin{bmatrix} e^{\alpha} & 0 \\ 0 & e^{\alpha} \end{bmatrix} \cdot \begin{bmatrix} \cos(\beta) & \sin(\beta) \\ -\sin(\beta) & \cos(\beta) \end{bmatrix} = \begin{bmatrix} e^{\alpha}\cos(\beta) & e^{\alpha}\sin(\beta) \\ -e^{\alpha}\sin(\beta) & e^{\alpha}\cos(\beta) \end{bmatrix}.$$

  In a similar way we can find that $e^{tJ} = \begin{bmatrix} e^{t\alpha}\cos(t\beta) & e^{t\alpha}\sin(t\beta) \\ -e^{t\alpha}\sin(t\beta) & e^{t\alpha}\cos(t\beta) \end{bmatrix}$.

- So how to use all of this knowledge? Well, let's give one example. Suppose we are given the 2-dimensional system $\begin{cases} x_1' = -2x_1, \\ x_2' = -x_1 - 2x_2, \end{cases}$ with initial values $x_1(0) = -1$, $x_2(0) = 3$. This is a linear system $x' = Ax$ with $A = \begin{bmatrix} -2 & 0 \\ -1 & -2 \end{bmatrix}$ and initial vector $x_0 = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$.

First we want to find the eigenvalues of A. We have $\det(A-\lambda I) = \det\begin{bmatrix} -2-\lambda & 0 \\ -1 & -2-\lambda \end{bmatrix} = (-2-\lambda)\,(-2-\lambda)-(-1)\cdot 0 = (\lambda+2)^2$, so A has a double eigenvalue $\lambda = -2$. So we are in case (ii) or (iii) of Section 1.4 of these notes. To find out which of the two cases we are, we need to find out how many linearly independent eigenvectors there are. Eigenvectors are found by looking for vertices v such that $Av = \lambda v$. Taking $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$, we get the two equations $\begin{cases} -2v_1 & = -2v_1. \\ -v_1 - 2v_2 & = -2v_2. \end{cases}$ This system gives as only information $v_1 = 0$, while v_2 can be anything. So all eigenvectors have the form $v = \begin{bmatrix} 0 \\ v_2 \end{bmatrix}$, and hence there is only one linearly independent eigenvector. We conclude that we are in case (iii).

According to Case (iii), this means that A has the Jordan form $J = \begin{bmatrix} -2 & 1 \\ 0 & -2 \end{bmatrix}$ and that there is an invertible real matrix P so that $A = PJP^{-1}$. To find P we can follow the recipe using eigenvectors from Section 1.4 above. But we can also use the knowledge that P exists. I.e. let's just try to find a P so that $A = PJP^{-1}$. This equation can be rewritten as $AP = PJ$. Filling in the entries for A and J, and setting $P = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the equation $AP = PJ$ corresponds to the matrix equation $\begin{bmatrix} -2a & -2b \\ -a-2c & -b-2d \end{bmatrix} = \begin{bmatrix} -2a & a-2b \\ -2c & c-2d \end{bmatrix}$. This matrix equation gives us four equations for the four unknowns a, b, c, d:

$$-2a \;=\; -2a, \quad -2b \;=\; a-2b, \quad -a-2c \;=\; -2c, \quad -b-2d \;=\; c-2d.$$

But there is in fact very little information we really get from these equations: $a = 0$ and $c = -b$, for all others we have a free choice. (This is in general the case, because the matrix P is never unique.) To keep life simple (remember, we also need to determine P^{-1}), we take $a = 0$, $b = 1$, $c = -1$, and $d = 0$, so $P = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. Then we have $P^{-1} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

Now we have all the knowledge we need to write out the solution of the systems of ODEs with initial values.

Since $J = \begin{bmatrix} -2 & 1 \\ 0 & -2 \end{bmatrix}$, from Case (iii) in this section we learn that $e^{tJ} = \begin{bmatrix} e^{-2t} & te^{-2t} \\ 0 & e^{-2t} \end{bmatrix}$. And then we can use the observations from page 1 of these extra notes to deduce

$$e^{tA} \;=\; P\begin{bmatrix} e^{-2t} & te^{-2t} \\ 0 & e^{-2t} \end{bmatrix}P^{-1} \;=\; \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} e^{-2t} & te^{-2t} \\ 0 & e^{-2t} \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \;=\; \begin{bmatrix} e^{-2t} & 0 \\ -te^{-2t} & e^{-2t} \end{bmatrix}.$$

Finally we use Theorem 1.5.5 from the Lecture Notes to conclude that the 2-dimensional system $x' = Ax$ with initial value $x(0) = x_0$ has the unique solution

$$x(t) \;=\; e^{tA}x_0 \;=\; \begin{bmatrix} e^{-2t} & 0 \\ -te^{-2t} & e^{-2t} \end{bmatrix}\begin{bmatrix} -1 \\ 3 \end{bmatrix} \;=\; \begin{bmatrix} -e^{-2t} \\ 3e^{-2t} + te^{-2t} \end{bmatrix}.$$

In other words, the solution is $\begin{cases} x_1(t) = -e^{-2t}, \\ x_2(t) = 3e^{-2t} + te^{-2t}. \end{cases}$

1.6 Decoupled systems

An n-dimensional system of differential equations $x' = ft, x)$ is called *decoupled* if we can partition the coordinates $\{1, \ldots, n\}$ into two parts J_1 and J_2 (so $J_1 \cup J_2 = \{1, \ldots, n\}$ and $J_1 \cap J_2 = \varnothing$), so that if $i \in J_1$ then the derivative x_i' depends on t and on x_j with $j \in J_1$ only, while if $i \in J_2$, then the derivative x_j' depends on t and those x_j with $j \in J_2$ only.

In other words, in a decoupled system we can divide the system of differential equations into two smaller systems that are completely independent from one another. In particular, the solutions and qualitative behaviour of the big system is completely determined by the solutions and qualitative behaviour of the two parts who don't interact with one another.

- An example of a decoupled system is the system $x' = Ax$ with $A = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$, which can be decoupled to the two equations $x_1' = \lambda_1 x_1$ and $x_2' = \lambda_2 x_2$.

- More explicitly, suppose we are given an n-dimensional system $x' = f(t, x)$, with $x = (x_1, \ldots, x_n)$, which can be decoupled as $\begin{cases} x_1' &= f_1(t, x_1, \ldots, x_m), \\ x_2' &= f_2(t, x_1, \ldots, x_m), \\ \vdots &= \vdots \\ x_m' &= f_m(t, x_1, \ldots, x_m), \\ x_{m+1}' &= f_{m+1}(t, x_{m+1}, \ldots, x_n), \\ x_{m+2}' &= f_{m+2}(t, x_{m+1}, \ldots, x_n), \\ \vdots &= \vdots \\ x_n' &= f_n(t, x_{m+1}, \ldots, x_n). \end{cases}$ Then we can consider the smaller systems $\begin{cases} x_1' &= f_1(t, x_1, \ldots, x_m), \\ \vdots &= \vdots \\ x_m' &= f_m(t, x_1, \ldots, x_m), \end{cases}$ and $\begin{cases} x_{m+1}' &= f_{m+1}(t, x_{m+1}, \ldots, x_n), \\ \vdots &= \vdots \\ x_n' &= f_n(t, x_{m+1}, \ldots, x_n), \end{cases}$ separately. From the first system we obtain the solutions for $x_1(t), \ldots, x_m(t)$, and from the second system the solutions for $x_{m+1}(t), \ldots, x_n(t)$. Then the full list $x(t) = (x_1(t), \ldots, x_m(t), x_{m+1}(t), \ldots, x_n(t))$ is a solution for the original large system.

- The partition into two parts is not always unique. For instance the system $\begin{cases} x_1' = x_1, \\ x_2' = 2x_2, \\ x_3' = -x_3, \end{cases}$ can be decoupled into $\begin{cases} x_1' = x_1, \\ x_2' = 2x_2, \end{cases}$ and $x_3' = -x_3$; but also in $x_1' = x_1$ and $\begin{cases} x_2' = 2x_2, \\ x_3' = -x_3. \end{cases}$

1.7 Higher dimensional linear systems

The analysis of linear systems of the form $x' = Ax$ of dimension $n > 2$ can be continued and works very similar to the 2-dimensional case. We won't do much about proving this, but just give the results. In order to do so, it makes sense to start with the 1-dimensional case.

- **Property 1.6**
  The only form of a 1-dimensional linear differential equation is $x' = Ax$, where $A = (\lambda)$ for some $\lambda \in \mathbb{R}$.

  The solution for this system is $x(t) = e^{\lambda t} x_0$ for a constant $x_0 \in \mathbb{R}$.

- We now formulate the results obtained in Sections 1.4 of these extra notes slightly differently.

  Property 1.7
  For every 2-dimensional linear system of differential equations $x' = Ax$ there exists an invertible real matrix P such that $A = PJP^{-1}$ where J has one of the following forms:

  (a) $J = \left(\begin{array}{c|c} B & 0 \\ \hline 0' & C \end{array}\right)$, *where B, C are real matrices of dimension smaller than 2 and $0, 0'$ are blocks of zeros of appropriate size;*

  (b) $J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$, *where $\lambda \in \mathbb{R}$;*

  (c) $J = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$, *where $\alpha, \beta \in \mathbb{R}$, $\beta > 0$.*

  The formulation with the sub-matrices in case (a) is a little overkill here, since $B, C, 0, 0'$ are just 1-dimensional matrices, hence just real numbers. But it prepares for the higher dimensional results below. And it also clearly indicates that we can consider case (a) as a decoupled system consisting of two smaller systems of lower dimension.

- For 3-dimensional systems we get the following:

  Property 1.8
  For every 3-dimensional linear system of differential equations $x' = Ax$ there exists an invertible real matrix P such that $A = PJP^{-1}$ where J has one of the following forms:

  (a) $J = \left(\begin{array}{c|c} B & 0 \\ \hline 0' & C \end{array}\right)$, *where B, C are real matrices of dimension smaller than 3 and $0, 0'$ are blocks of zeros of appropriate size;*

  (b) $J = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$, *where $\lambda \in \mathbb{R}$.*

  If you followed what was the relation between eigenvalues and the Jordan forms of 2-dimensional matrices, then you should have some idea when the Jordan form in (b) appears. That is if A has one triple eigenvalue λ with only one corresponding eigenvector.

  In case (a) above, we can assume that C is 1-dimensional, hence just a real number, and B is one of the 2-dimensional cases from the previous property. In particular, in (a) we have a decoupled system, whereas the system in (b) cannot be decoupled. More explicitly, if we fill in the different possibilities for B and C in case (a) we get the following options:

  (i) $J = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$, where $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$;

  (ii) $J = \begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}$, where $\lambda_1, \lambda_2 \in \mathbb{R}$;

  (iii) $J = \begin{pmatrix} \alpha & \beta & 0 \\ -\beta & \alpha & 0 \\ 0 & 0 & \lambda \end{pmatrix}$, where $\alpha, \beta, \lambda \in \mathbb{R}$, $\beta > 0$.

  Together with the special form in (b) this means we have four different Jordan forms for a real 3×3 matrix.

- For 4-dimensional system, a new type appears, as is formulated in the following result.

  Property 1.9
  For every 4-dimensional linear system of differential equations $x' = Ax$ there exists an invertible real matrix P such that $A = PJP^{-1}$ where J has one of the following forms:

  (a) $J = \left(\begin{array}{c|c} B & 0 \\ \hline 0' & C \end{array}\right)$, *where B, C are real matrices of dimension smaller than 4 and $0, 0'$ are blocks of zeros of appropriate size;*

  (b) $J = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}$, *where $\lambda \in \mathbb{R}$;*

  (c) $J = \begin{pmatrix} \alpha & \beta & 1 & 0 \\ -\beta & \alpha & 0 & 1 \\ 0 & 0 & \alpha & \beta \\ 0 & 0 & -\beta & \alpha \end{pmatrix}$, *where $\alpha, \beta \in \mathbb{R}$, $\beta > 0$.*

  Again, the special Jordan form in (b) appears if A has a 4-fold eigenvalue λ with only one corresponding eigenvector.

  The form in (c) is new. That form corresponds to the following special matrices A: A has a double complex eigenvalue $\alpha + \beta i$ with only one corresponding eigenvector. Then also the conjugate complex number $\alpha - \beta i$ is a double eigenvalue with only one corresponding eigenvector. Because of the discrepancy between the algebraic multiplicity two of the eigenvalues and the fact that they have only one eigenvector, we get that extra block with 0s and 1s in the top right corner of the Jordan form.

  The smaller matrices B, C in part (a) can either both be 2-dimensional, and then according to the property for 2-dimensional systems, or C is just 1-dimensional (just a real number) and B is 3-dimensional and one of the special types for that dimension.

- When we go to even higher dimensions, no new forms appear.

  Property 1.10
  For every n-dimensional linear system of differential equations $x' = Ax$ there exists an invertible real matrix P such that $A = PJP^{-1}$ where J has one of the following forms:

  (a) $J = \left(\begin{array}{c|c} B & 0 \\ \hline 0' & C \end{array}\right)$, *where B, C are real matrices of dimension smaller than $n$and $0, 0'$ are blocks of zeros of appropriate size;*

  (b) $J = \begin{pmatrix} \lambda & 1 & 0 & \cdots & \cdots & 0 \\ 0 & \lambda & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 1 & 0 \\ 0 & \cdots & \cdots & 0 & \lambda & 1 \\ 0 & \cdots & \cdots & \cdots & 0 & \lambda \end{pmatrix}$, *where $\lambda \in \mathbb{R}$, and we must have dimension $n \geq 2$;*

$$\text{(c)}\quad J = \begin{pmatrix}
\alpha & \beta & 1 & 0 & 0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 \\
-\beta & \alpha & 0 & 1 & 0 & 0 & \ddots & & & & 0 & 0 \\
0 & 0 & \alpha & \beta & 1 & 0 & 0 & \ddots & & & \vdots & \vdots \\
0 & 0 & -\beta & \alpha & 0 & 1 & 0 & \ddots & \ddots & & \vdots & \vdots \\
\vdots & \vdots & 0 & 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\
\vdots & \vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 & \vdots \\
\vdots & \vdots & & & \ddots & \ddots & \ddots & \ddots & 1 & 0 & 0 & 0 \\
\vdots & \vdots & & & & \ddots & \ddots & \ddots & 0 & 1 & 0 & 0 \\
0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 & \alpha & \beta & 1 & 0 \\
0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 & -\beta & \alpha & 0 & 1 \\
0 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 0 & \alpha & \beta \\
0 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 0 & -\beta & \alpha
\end{pmatrix},$$

where $\alpha, \beta \in \mathbb{R}$, $\beta > 0$, and we must have that the dimension n is even.

It is possible to give explicit solutions for (b) and (c), and for (a) a solution is obtained by combining the solutions for the two decoupled parts. But since we're not really that much interested in the solutions, but more in the qualitative behaviour, we won't discuss these.

Exercises

1 For each matrix A given below, find its Jordan form J and find the matrix P so that $A = PJP^{-1}$.

$$\text{(a)}\quad A = \begin{bmatrix} 0 & 1 \\ -2 & 3 \end{bmatrix}; \qquad \text{(b)}\quad A = \begin{bmatrix} 41 & -29 \\ 58 & -41 \end{bmatrix}; \qquad \text{(c)}\quad A = \begin{bmatrix} 9 & 4 \\ -9 & -3 \end{bmatrix}.$$

2 For each matrix A in question 1 above, calculate e^{tA}.

3 Show that the two statements in Theorem 1.2 on page 3 of these notes are not true in general if we allow the matrix A to have complex entries.

4 Prove Theorem 1.4 on page 4 of these notes. I.e. show that the eigenvectors of different eigenvalues are linearly independent.

5 Prove part (b) of Theorem 1.5 on page 4.

6 Prove that if A is an $n \times n$ matrix, and P is an $n \times n$ invertible matrix, then A and $P^{-1}AP$ have the same eigenvalues.

7 Prove the final statement of Section 1.4 on page 6 of these notes: *If 2×2 matrices A, B are similar, then there exist invertible matrices P_A, P_B so that the Jordan forms $J_A = P_A^{-1}AP_A$ and $J_B = P_B^{-1}BP_B$ are the same.*