Error control in scientific modelling (MATH 500, Herbst)

Sheet 8: Matrix perturbation theory

```
# Install some packages
begin
    using LinearAlgebra
    using PlutoUI
    using PlutoTeachingTools
    using Plots
end
```

Exercise 1

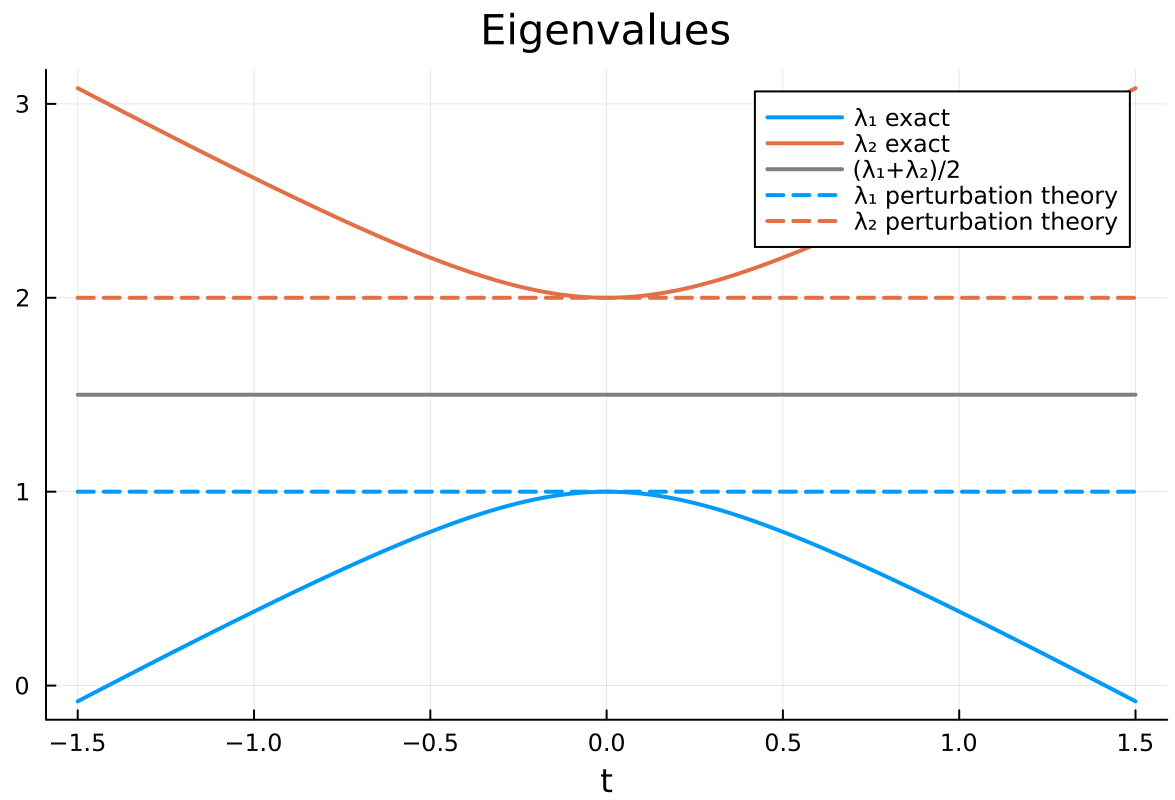
In this exercise we will compute some of the results of the lecture explicitly for the problem $A(t) = \tilde{A} + t\Delta A$ with

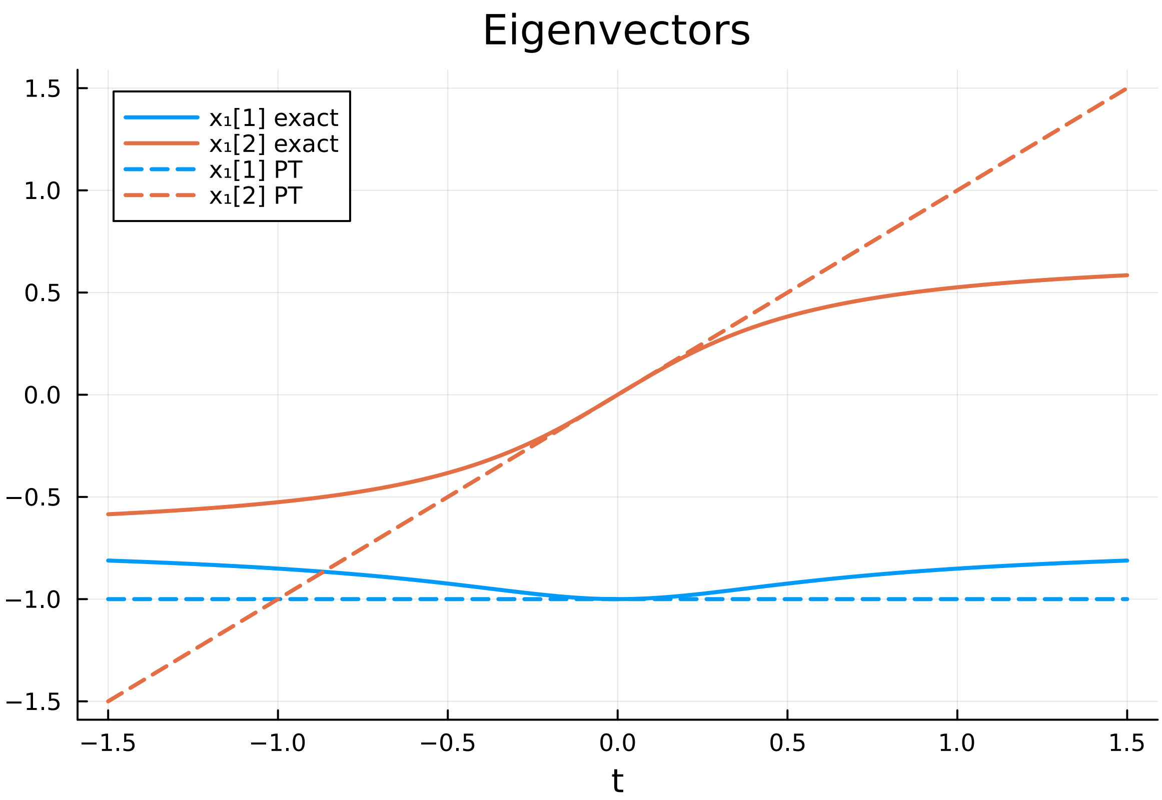
$$\tilde{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1+\delta \end{pmatrix} \qquad \Delta A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

For notational simplicity let $\lambda_1(t)$ denote the exact first eigenvalue of $A(t)$ and $\lambda_2(t)$ denote the exact second eigenvalue of $A(t)$, ordered by size.

```
begin
    Atilde = Diagonal([1, 1 + δ])  # Reference matrix
    ΔA = [0 1; 1 0]                # Perturbation
    A(t) = Atilde + t * ΔA         # Matrix depending on t
    ts = -1.5:0.025:1.5            # Range of values for t for plotting
end;
```

Gap δ = 1.0


Eigenvalues
λ₁ exact
λ₂ exact
(λ₁+λ₂)/2
λ₁ perturbation theory
λ₂ perturbation theory
3
2
1
0
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
t




```julia
let
    p = plot(title="Eigenvectors")
    xlabel!(p, "t")
    if !ismissing(x1_exact)
        plot!(p, ts, getindex.(x1_exact, Ref(1)); label="x₁[1] exact", lw=2, c=1)
        plot!(p, ts, getindex.(x1_exact, Ref(2)); label="x₁[2] exact", lw=2, c=2)
    end

    if !ismissing(Δx1)
        i_t0 = findfirst(iszero, ts)
        x_tilde = x1_exact[i_t0]

        x1_PT = [x_tilde + t * Δx1 for t in ts]
        plot!(p, ts, getindex.(x1_PT, Ref(1)); label="x₁[1] PT", lw=2, c=1, ls=:dash)
        plot!(p, ts, getindex.(x1_PT, Ref(2)); label="x₁[2] PT", lw=2, c=2, ls=:dash)
    end

    p
end
```

Eigenvalues. First we are concerned with exploring the exact eigenvalues. Code up a function, that computes the *exact* first and second eigenvalue of $A(t)$ at the t-values given in `ts` and store the results in `λ1_exact` and `λ2_exact`.

```julia
begin
    λ_exact = [sort(eigvals(A(t)), by=abs) for t in ts]
    λ1_exact = getindex.(λ_exact, Ref(1))
    λ2_exact = getindex.(λ_exact, Ref(2))
end;
```

In the plot we clearly see the linear dependency in t for large values of $|t|$ predicted by Proposition 7.

Now vary the slider of δ. You should notice that the dependency of the eigenvalues of $A(t)$ in t is smooth as long as the eigvalues of $\tilde{A}$ are non-degenerate (i.e. $\delta > 0$). However, for $\delta = 0$ we observe a clear kink.

Now we want to check Theorem 6, which suggests that the *average* eigenvalue

$$\frac{1}{2}(\lambda_1(t) + \lambda_2(t))$$

should be smooth for all t.

Compute the average eigenvalue for each value of t:

```
λ_exact_avg = (λ1_exact .+ λ2_exact) ./ 2;
```

You should observe in the plot that this average is continuous for all values of δ.

Now compute the *first-order perturbative correction* $\Delta\lambda_1 \equiv \lambda_1'(0)$ and $\Delta\lambda_2 \equiv \lambda_2'(0)$ according to the perturbation theory treatment we proposed in the lecture, code them up here:

```
begin
    Δλ_1 = [1; 0]' * ΔA * [1; 0]
    Δλ_2 = [0; 1]' * ΔA * [0; 1]
end;
```

From this we can easily compute the first order (in t) approximation of $\lambda_1(t)$ and $\lambda_2(t)$ as the sum $\tilde{\lambda}_1 + t\,\Delta\lambda_1$, where $\tilde{\lambda}_1$ is the first eigenvalue of $\tilde{A}$. Code this up below:

```
begin
    λ1_tilde = 1.0
    λ2_tilde = 1.0 + δ

    λ1_PT = λ1_tilde .+ ts .* Δλ_1
    λ2_PT = λ2_tilde .+ ts .* Δλ_2
end;
```

Eigenvectors. Next we are concerned with the eigenvectors. Code up the computation of the *exact* eigenvector $x_1(t)$ of $A(t)$ corresponding to $\lambda_1(t)$ at all t in `ts`. Make sure to produce a list of vectors.

Make sure you use a homogeneous sign convention for all t as otherwise the plot generated in the "Eigenvectors" graph above may not be continuous (as it should be).

```
x1_exact =
  [[-0.811242, -0.58471], [-0.812724, -0.582649], [-0.814245, -0.580521], [-0.815809, -0.57
```

```julia
x1_exact = [eigen(Hermitian(A(t))).vectors[:, 1] for t in ts]
```

Next up is the first-order correction $\Delta x_1 = x_1'(0)$ from perturbation theory. For the computation keep in mind that $(\tilde{A} - \tilde{\lambda} I)$ is singular, but that the eigenvector corresponding to the 0-eigenvalue is actually projected out by $Q_{\tilde{\lambda}}$, such that you can arbitrarily shift the eigenvalue along this mode to avoid issues when computing the inverse.

```
[-0.0, 1.0]
```

```julia
begin
    i_t0 = findfirst(iszero, ts)
    x_tilde = x1_exact[i_t0]

    Q = [0 0;
         0 1]
    A_shifted = (Atilde - λ1_tilde * I)
    A_shifted[1, 1] = 1e-12

    Δx1 = - A_shifted \ (Q * ΔA * x_tilde)
end
```

Exercise 2

Above we constructed the perturbative expansion of the eigenpairs of $A(t) = \tilde{A} + t\Delta A$ starting from $\tilde{A}$ and its eigenpairs $(\tilde{\lambda}_i, \tilde{x}_i)$ as the reference. We computed corrections to first order, i.e. obtained an expansion

$$
\begin{aligned}
\lambda_1(t) &= \tilde{\lambda}_1 + t\Delta\lambda_1 + O(|t|^2) \\
x_1(t) &= \tilde{x}_1 + t\Delta x_1 + O(|t|^2).
\end{aligned}
$$

In this exercise we want to compute a lower bound for the radius of convergence for this expansions using Theorem 6.4.

As a reminder this Theorem guarantees that $\lambda_1(t)$ and $x_1(t)$ are analytic as long as $|t| < \varrho_a$ with

$$
\varrho_a = \inf_{z \in \Gamma} \Big[\varrho\big(T(z)\big) \Big]^{-1} \quad \text{with} \quad T(z) = R_z(\tilde{A}) \Delta A
$$

and Γ a contour in the complex plane enclosing only $\tilde{\lambda}_1$ and $\varrho\,(T(z))$ denoting the spectral radius of $T(z)$. As a result the above expansions are guaranteed to converge for $|t| < \varrho_a$ [1].

We consider the same case as above, i.e.

$$\tilde{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1+\delta \end{pmatrix} \quad \text{and} \quad \Delta A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

but for simplicity we will assume $\delta > 0$.

(a) Show that the spectral radius of $R_z(\tilde{A})\Delta A$ is

$$\varrho\left(R_z(\tilde{A})\Delta A\right) = \frac{1}{\sqrt{|1-z|}} \frac{1}{\sqrt{|1+\delta-z|}}$$

(a) solution We note

$$R_z(\tilde{A}) = \begin{pmatrix} \frac{1}{1-z} & 0 \\ 0 & \frac{1}{1+\delta-z} \end{pmatrix} \quad \text{and} \quad T(z) = R_z(\tilde{A})\,\Delta A = \begin{pmatrix} 0 & \frac{1}{1-z} \\ \frac{1}{1+\delta-z} & 0 \end{pmatrix}.$$

We proceed to compute the eigenvalues of $T(z)$ for which we note that its trace (sum of eigenvalues) is zero and its determinant (product of eigenvalues) is

$$\det T(z) = -\frac{1}{(1-z)(1+\delta-z)},$$

from which we deduce its eigenvalues to be

$$\pm\frac{1}{\sqrt{(1-z)(1+\delta-z)}}$$

Therefore its spectral radius is

$$\varrho(T(z)) = \left|\frac{1}{\sqrt{(1-z)(1+\delta-z)}}\right| = \frac{1}{|\sqrt{1-z}|} \frac{1}{|\sqrt{1+\delta-z}|} = \frac{1}{\sqrt{|1-z|}} \frac{1}{\sqrt{|1+\delta-z|}}$$

(b) Next we want to compute ϱ_a. For this we take a circular contour enclosing only $\lambda_1 = 1$, i.e. the contour is parametrised by

$$\Gamma = \left\{1 + re^{i\theta} \mid \theta \in [0, 2\pi)\right\}$$

where $0 < r < \delta$ to avoid enclosing $\lambda_2 = 1 + \delta$. With this ansatz

$$\varrho_a = \inf_{z\in\Gamma} \left[\varrho(T(z))\right]^{-1} = \inf_{\theta\in[0,2\pi]} \left[\varrho\left(T\left(1+re^{i\theta}\right)\right)\right]^{-1}.$$

Show that $\varrho_a = \sqrt{r\delta - r^2}$.

(b) solution First we consider the argument of the infimum:

$$
\begin{aligned}
\left[\varrho(T(1+re^{i\theta}))\right]^{-1} &= \sqrt{|1-(1+re^{i\theta})|}\,\sqrt{|1+\delta-(1+re^{i\theta})|} \\
&= \sqrt{r}\,\sqrt{|1-re^{i\theta}|} \\
&= \sqrt{r}\,\sqrt[4]{[\delta-r\cos(\theta)]^2+r^2\sin^2(\theta)} \\
&= \sqrt{r}\,\sqrt[4]{\delta^2-2\delta r\cos(\theta)+r^2}
\end{aligned}
$$

Since $r > 0$ and $\delta > 0$ this expression is smallest when $\cos(\theta)$ is maximal, i.e. 1. Therefore

$$
\varrho_a = \sqrt{r}\,\sqrt[4]{\delta^2-2\delta r+r^2} = \sqrt{r}\,\sqrt[4]{(\delta-r)^2} = \sqrt{r\delta-r^2}
$$

(c) The final step is to optimise the contour itself, such that ϱ_a (and thus the possible values for $|t|$) are largest. Show that we can guarantee convergence of the perturbation expansion for $|t| < \delta/2$.

(c) solution Since $f(r) = r\delta - r^2$ has a maximum when

$$
0 = f'(r) = \delta - 2r,
$$

the best choice is $r = \delta/2$, yielding

$$
\varrho_a = \sqrt{\delta^2/2 - \delta^2/4} = \sqrt{\delta^2/4} = \delta/2.
$$

[1]:

> Note that this is an underestimation, the true radius of convergence depends on the analyticity of $\int_\Gamma R_z(A(t))\mathrm{d}z$.