A First Stab at the Big ODE Theorems

Suppose $S$ is an $n \times n$ real and symmetric matrix, that is $S \in \mathbb{R}^{n \times n}$, $S = S^T$. You should know that such matrices always have a basis of real eigenvectors with corresponding real eigenvalues, that is

$$\text{span}\{r_1, \ldots, r_n\} = \mathbb{R}^n, \quad Sr_k = \lambda_k r_k \text{ where } \lambda_k \in \mathbb{R}.$$ 

Moreover, since

$$Sr_i = \lambda_i r_i \implies r_j^T Sr_i = \lambda_i r_j^T r_i$$
$$Sr_j = \lambda_j r_j \implies r_i^T Sr_j = \lambda_j r_i^T r_j$$

and from the fact that $S$ is symmetric

$$r_j^T Sr_i = (r_j^T Sr_i)^T = r_i^T S^T r_j = r_i^T Sr_j$$

allows us to conclude

$$(\lambda_i - \lambda_j) r_j^T r_i = r_j^T Sr_i - r_i^T Sr_j = 0.$$ 

This shows that when $\lambda_i \neq \lambda_j$ we must have $r_j \cdot r_i = r_j^T r_i = 0$. That is, for a symmetric matrix, eigenvectors associated to different eigenvalues are necessarily orthogonal (perpendicular).

With this in mind, consider the following symmetric matrix

$$S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \implies \det(S - \lambda I) = (1 - \lambda)^2(2 - \lambda) = 0 \implies \lambda = 1 \text{ or } \lambda = 2.$$ 

The eigenvalue $\lambda = 2$ has eigenvector $r_{\lambda=2} = (0 \ 0 \ 1)^T$, whereas any nonzero vector of the form $r_{\lambda=1} = (\alpha \ \beta \ 0)^T$ is an eigenvector associated to eigenvalue $\lambda = 1$. Notice that the eigenspace associated to eigenvalue $\lambda = 2$ is one dimensional. However, the eigenspace associated to eigenvalue $\lambda = 1$ is two dimensional. A basis to this two dimensional eigenspace need not be orthogonal, e.g. $\{(1 \ 0 \ 0)^T, \ (0 \ 1 \ 0)^T\}$, but it is always possible to build one that is, e.g. $\{(1 \ 0 \ 0)^T, \ (0 \ 1 \ 0)^T\}$.

This example together with the preceding paragraph illustrates a very important fact. Not only do real $n \times n$ symmetric matrices have all real eigenvalues and a basis of eigenvectors $\mathbb{R}^n = \text{span}\{r_1, \ldots, r_n\}$, these eigenvectors can be designed so that

$$r_i \cdot r_j = \delta_{i,j}, \quad \text{where } \delta_{i,j} = \begin{cases} 0 & \text{when } i \neq j \\ 1 & \text{when } i = j. \end{cases}$$

That is, every symmetric matrix has an orthonormal basis of eigenvectors.
1. Find an orthonormal basis of eigenvectors for the symmetric matrix \(
\begin{pmatrix}
3 & -1 \\
-1 & 3
\end{pmatrix}
\).

2. Find an orthonormal basis of eigenvectors for the symmetric matrix \(
\begin{pmatrix}
4 & 0 & 2 \\
0 & 6 & 0 \\
2 & 0 & 4
\end{pmatrix}
\).

3. Find an orthonormal basis of eigenvectors for the symmetric matrix \(
\begin{pmatrix}
4 & 1 & 1 \\
1 & 4 & 1 \\
1 & 1 & 4
\end{pmatrix}
\).

Hint: This matrix has two eigenvalues, \(\lambda = 3\) with multiplicity two, and \(\lambda = 6\) with multiplicity one.

4. Suppose \(S\) is real symmetric. Conclude there is an orthogonal matrix \(O\) (an orthogonal matrix satisfies \(O^T = O^{-1}\)) such that \(\Lambda = O^T S O\) where \(\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)\).

5. Apply the result in exercise 4 to the matrices given in exercises 1, 2 and 3. Use the orthonormal basis you determined to form \(O\).

---

The vector norm you all saw in calculus is the *Euclidean norm*

\[ ||\mathbf{x}|| \equiv \sqrt{x_1^2 + \cdots + x_n^2}. \]

You probably have not seen the associated matrix norm

\[ |||A||| \equiv \max_{||\mathbf{x}|| = 1} ||Ax||. \]

This really is a norm on the vector space of \(n \times n\) real matrices. In an exercise below you are asked to deduce the explicit formula

\[ |||A||| = \sqrt{\max(\lambda_1, \ldots, \lambda_n)} \text{ where the } \lambda_k \text{'s denote the eigenvalues of } A^T A. \]

One very important property of the matrix norm so defined is:

\[ \forall \mathbf{x} \in \mathbb{R}^n \text{ we have } ||Ax|| \leq |||A||| ||\mathbf{x}||. \]

This can be shown as follows.

\[ ||Ax||^2 = (Ax)^T (Ax) = x^T A^T A x. \]

The matrix \(A^T A\) is symmetric and can therefore be diagonalized by an orthogonal matrix

\[ O^T A^T A O = \Lambda \quad \Rightarrow \quad A^T A = O \Lambda O^T. \]

Clearly, each entry in the diagonal matrix of eigenvalues is nonnegative since

\[ ||Ax||^2 = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2, \text{ where } y = O^T \mathbf{x}, \text{ (or } \mathbf{x} = O y), \]

2
and furthermore since \( \|y\|^2 = y^T y = x^T O O^T x = \|x\|^2 \) and
\[
\lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 \leq \max(\lambda_1, \ldots, \lambda_n) (y_1^2 + \cdots + y_n^2) = \|A\|^2 \|y\|^2,
\]
we deduce the fact \( \|Ax\| \leq \|A\| \|x\| \).

Let \( \sigma^2(A) \) denote the largest eigenvalue of \( A^T A \). That is, \( \|A\| = \sqrt{\sigma^2(A)} \). It is worth noting that \( \sigma^2(A) \) is not a particularly easy number to compute. As you might know however, the trace of a matrix is invariant under similarity transformation. Therefore,
\[
\text{Tr}(A^T A) = \lambda_1 + \cdots + \lambda_n \geq \max(\lambda_1, \ldots, \lambda_n)
\]
\[
\Rightarrow \|A\|^2 = \sigma^2(A) \leq \text{Tr}(A^T A) = \sum_{i=1}^n \sum_{j=1}^n a_{i,j}^2 \equiv \|A\|_F^2.
\]

\( \|A\|_F \) is called the Frobenius norm of \( A \), a much easier quantity to compute. So we have also shown
\[
\forall x \in \mathbb{R}^n \text{ we have } \|Ax\| \leq \|A\|_F \|x\|.
\]

Here’s an example. The matrix norm of \( A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \) is computed as follows.
\[
A^T A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \Rightarrow \lambda = \frac{3 \pm \sqrt{5}}{2} \Rightarrow \|A\| = \sqrt{\frac{3 + \sqrt{5}}{2}}.
\]
The Frobenius norm is much easier to compute.
\[
\|A\|_F^2 = 1^2 + 1^2 + 0^2 + 1^2 \Rightarrow \|A\|_F = \sqrt{3}.
\]

6. Recall \( \sigma^2(A) \) was defined to be the largest eigenvalue of \( A^T A \), and we showed that \( \|Ax\|^2 \leq \sigma^2(A) \|x\|^2 \) for every vector \( x \in \mathbb{R}^n \). You show there is a particular vector \( x_* \) satisfying \( \|x_*\| = 1 \) such that \( \|Ax_*\|^2 = \sigma^2(A) \). Therefore, conclude \( \|A\| = \sqrt{\sigma^2(A)} \).

Hint: Take \( x_* \) to be an appropriate eigenvector to \( A^T A \).

7. Compute the matrix norm and the Frobenius norm of \( A = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} \).

8. Compute the matrix norm and the Frobenius norm of \( A = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{pmatrix} \).

One more topic to go.
The Grönwall inequality reads as follows. Suppose for a function of $t$, say $w(t)$, we know
\[ w(t) \leq \alpha + \beta \int_{\min(t,t_0)}^{\max(t,t_0)} w(\tau) \, d\tau + \gamma |t - t_0| \]
for constants $\alpha$, $\beta > 0$ and $\gamma$. Then, we must have
\[ w(t) \leq \alpha e^{\beta |t-t_0|} + \gamma (e^{\beta |t-t_0|} - 1)/\beta. \]

Before deriving Grönwall, let me give an application. Consider two scalar IVPs
\[ \frac{du}{dt} = f(u) \quad \text{and} \quad \frac{dv}{dt} = f(v) \]
with initial conditions $u(0) = u_0$ and $v(0) = v_0$. The functions $u$ and $v$ solve the same differential equation but with different initial data. Notice that
\[ \frac{du}{dt} - \frac{dv}{dt} = f(u) - f(v) \quad \Rightarrow \quad u(t) - v(t) = u_0 - v_0 + \int_0^t (f(u(\tau)) - f(v(\tau))) \, d\tau. \]
Using the triangle inequality, together with Lipschitz continuity, we find
\[ |u(t) - v(t)| \leq |u_0 - v_0| + L \int_{\min(t,0)}^{\max(t,0)} |u(\tau) - v(\tau)| \, d\tau. \]
Apply Grönwall’s inequality to conclude
\[ |u(t) - v(t)| \leq |u_0 - v_0| e^{L|t|}. \]
But this result tells us a lot. First, it says the solution to the IVP must be unique. When $u_0 = v_0$ we must have $u(t) = v(t)$ for all $t$. Second, for fixed $t$, the solution to this IVP depends continuously on its initial data. That is, $u(t)$ changes only by a little bit if its initial data $u_0$ is changed by a little bit.

For the last two exercises, consider the $n$ dimensional, first order, linear, variable coefficient system
\[ \frac{d}{dt} u + A(t) u = b(t), \quad u(0) = u_0, \]
where you may assume for all $t$ that $||A(t)|| \leq a$ and $\|b(t)\| \leq b$. ($a$ and $b$ are constants.)

Use Grönwall’s inequality to prove the following.

9. Prove that $||u(t)||$ is bounded for all bounded $t$. That is, $u(t)$ can not blow-up in finite time. Hint: Show $||u(t)|| \leq ||u_0|| e^{a|t|} + b(e^{a|t|} - 1)/a.$

10. Prove the solution to the IVP is unique. Hint: As done in the application above, assume there are two solutions, say $u(t)$ and $v(t)$, and show $||u(t) - v(t)|| \leq 0.$
Here’s the derivation of Grönwall’s inequality. First, let me assume that \( t_0 = 0 \) and \( t \geq 0 \) and show
\[
w(t) \leq \alpha + \beta \int_0^t w(\tau) \, d\tau + \gamma t
\]
\[
\Rightarrow w(t) \leq \alpha e^{\beta t} + \gamma (e^{\beta t} - 1)/\beta.
\]
The other cases, i.e. \( t < 0 \) etc., can easily be deduced from this case by substitution.

Define \( W(t) = \int_0^t w(\tau) \, d\tau \), and use the Fundamental Theorem of Calculus to write the given inequality as
\[
\frac{dW}{dt} - \beta W \leq \alpha + \gamma t \quad \Rightarrow \quad \frac{d}{dt} (e^{-\beta t} W) \leq e^{-\beta t} (\alpha + \gamma t).
\]
Integrate from 0 to \( t \geq 0 \) to obtain
\[
W(t) \leq \frac{\alpha}{\beta} (e^{\beta t} - 1) - \frac{\gamma}{\beta} t + \frac{\gamma}{\beta^2} (e^{\beta t} - 1).
\]
Again, use the given inequality, substituting \( W(t) \) in for the integral, to finally get
\[
w(t) \leq \alpha + \left( \beta W(t) \right) + \gamma t \leq \alpha + \left( \alpha (e^{\beta t} - 1) - \gamma t + \frac{\gamma}{\beta} (e^{\beta t} - 1) \right) + \gamma t
\]
\[
= \alpha e^{\beta t} + \gamma (e^{\beta t} - 1)/\beta.
\]
Second, I’ll do the case when \( t_0 = 0 \) and \( t < 0 \) but will leave the \( t_0 \neq 0 \) case to you. Let \( t = -t' \) where \( t' > 0 \). So here we have to show
\[
w(-t') \leq \alpha + \beta \int_{-t'}^0 w(\tau) \, d\tau + \gamma t'
\]
\[
\Rightarrow w(-t') \leq \alpha e^{\beta t'} + \gamma (e^{\beta t'} - 1)/\beta.
\]
But this follows immediately from the case already treated by changing variable \( \tau = -t' \) in the integral above and calling \( \bar{w}(t) = w(-t) \).