oo Now what about the case of degenerate eigenvalues, i.e. eigenvalues w/ algebraic multiplicity > 1? First of all, if we are considering normal matrices, then we should still expect a set of eigenvectors that span the space. Where it gets odd is that for distinct eigenvalues the eigenvectors are unique, while for degenerate ones we have some freedom to choose them. This is because in the distinct case we are looking for a basis in 1D (which is unique when normalized), whereas in the degenerate case we are looking for a basis in $n \geq 2$ (which is obviously not unique).

If the original matrix A_0 has degenerate eigenvalues, then perturbing it may very well change this (for the entire matrix A may not have them), though at times this is not the case. If the perturbed matrix has distinct eigenvalues, then there is a unique basis in which this should be described. So we start out w/ the freedom to pick a basis, but may be forced into a particular one (which we do not know at the start).

So degenerate perturbation theory:

Suppose that A_0 has a degenerate eigenvalue $\lambda_n^{(0)}$ w/ n associated eigenvectors $y_{n,i}^{(0)}$ $(i = 1, 2, \cdots, n)$.

Then we know that $A_0 y_{n,i}^{(0)} = \lambda_n^{(0)} y_{n,i}^{(0)}$ and our expansions of λ and y (eigens of A):

$\lambda_{n,i} = \lambda_n^{(0)} + \epsilon \lambda_{n,i}^{(1)} + \epsilon^2 \lambda_{n,i}^{(2)} + \cdots$

$y_{n,i} = y_{n,i}^{(0)} + \epsilon y_{n,i}^{(1)} + \epsilon^2 y_{n,i}^{(2)} + \cdots$

Note no i label on $\lambda_n^{(0)}$ since it is degenerate, but on everything else (assuming ϵ breaks the degeneracy and $y_{n,i}$'s are distinct).

The goal is to pick the $y_{n,i}$'s that are compatible with the corrections brought by ϵ.

To start, we take the expansions above and plug them into $Ax = (A_0 + \epsilon A_1)x = \lambda x$ and rearranging/labelling just like last time obtain:

i=0 $\quad [A_0 - \lambda_n^{(0)}]\, y_{n,i}^{(0)} = 0$

i=1 $\quad [A_0 - \lambda_n^{(0)}]\, y_{n,i}^{(1)} = -[A_1 - \lambda_{n,i}^{(1)}]\, y_{n,i}^{(0)}$

i=2 $\quad [A_0 - \lambda_{n,i}^{(0)}]\, y_{n,i}^{(2)} = -[A_1 - \lambda_{n,i}^{(1)}]\, y_{n,i}^{(1)} + \lambda_{n,i}^{(2)} y_{n,i}^{(0)}$

i=3 $\quad [A_0 - \lambda_{n,i}^{(0)}]\, y_{n,i}^{(3)} = -[A_1 - \lambda_{n,i}^{(1)}]\, y_{n,i}^{(2)} + \lambda_{n,i}^{(2)} y_{n,i}^{(1)} + \lambda_{n,i}^{(3)} y_{n,i}^{(0)}$

Now remember that $y^{(0)}_{n,i}$ are unknown in the sense that they are undetermined by the expression $[A_0 - \lambda^{(0)}_n] y^{(0)}_{n,i} = 0$. So what we can do is just pick a basis of eigenvectors of A_0 associated w/ $\lambda^{(0)}_n$, call them $x^{(0)}_{n,i}$, as a starting point, and then define $y^{(0)}_{n,i} = \sum_{j=1}^{m} a_{ij} x^{(0)}_{n,j}$ as what will work w/ the ϵ corrections.
So we pick any orthonormal basis $x^{(0)}_{n,i}$ and let $\{a_{ij}\}$ be the unknowns to look for.

\ First of all we expect $y^{(0)}_{n,i}$ to be orthogonal, and normalizing them means
$(y^{(0)}_{n,i}, y^{(0)}_{n,k}) = \delta_{ik} \Rightarrow \sum_j a^*_{ij} a_{kj} = \delta_{ik}$.

Now recall that in order for eqns. 0-3 to have solutions, the r.h.s. of each must be orthogonal to whatever satisfies $[A^+ - \lambda^*_{n,i} I]\psi_i = 0$, but in this case, for self-adjoint $A = A^+$, λ is real so $\lambda = \lambda^* \Rightarrow [A - \lambda_{n,i} I]\psi_i = 0$ then the ψ_i's are all of the vectors associated w/ $\lambda_{n,i}$.

Using our starting point $\{x^{(0)}_{n,i}\}$ we can demand:
$(x^{(0)}_{n,j}, [A_1 - \lambda^{(1)}_{n,i}] y^{(0)}_{n,i}) = 0$ for all i and j, i.e. the r.h.s. is $\perp$ to the space spanned by $x^{(0)}_{n,j}$
but this becomes
$$\sum_k (x^{(0)}_{n,j}, [A_1 - \lambda^{(1)}_{n,i}] a_{ik} x^{(0)}_{n,k}) = \sum_k (x^{(0)}_{n,j}, [A_1 - \lambda^{(1)}_{n,i}] x^{(0)}_{n,k}) a_{ik} = 0$$
or
$$\sum_k [(x^{(0)}_{n,j}, A_1 x_{n,k}) - \lambda^{(1)}_{n,i} \delta_{jk}] a_{ik} = 0$$

Squinting your eyes, you might realize that this is an eigenvalue problem whose eigenvectors a_{ik}, $\{k = 1, .., m\}$ are the coefficients which will let us know which basis will work, i.e. $y^{(0)}_{n,i} = \sum_k a_{ik} x^{(0)}_{n,k}$.

In fact defining M w/ $M_{jk} = \underbrace{(x^{(0)}_{n,j}, A_1 x^{(0)}_{n,k})}$, then hiding the component label on a_{ik}, we have:
$M_{kj} = (x^{(0)}_{n,k}, A_1 x^{(0)}_{n,j}) = (A_1 x^{(0)}_{n,k}, x^{(0)}_{n,j}) = (x^{(0)}_{n,j}, A_1 x^{(0)}_{n,k})^*$
$(M - \lambda^{(1)}_{n,i}) a_i = 0$ $\quad$ A_1-Hermitian $\Rightarrow M_{jk} = M^*_{kj}$ - Hermitian
This is a Hermitian problem so we know $\lambda^{(1)}_{n,i}$ are real and we can take $(a_i, a_j) = \delta_{ij}$.
Solving the eigenvalue problem: (1) gives first order corrections to $\lambda^{(0)}_n$, i.e. $\lambda^{(1)}_n$, (2) determines correct linear combinations of $x^{(0)}_{n,i}$ via a_{ik}, (3) guarantees a solution to $[A_0 - \lambda^{(0)}_n] y^{(1)}_{n,i} = -[A_1 - \lambda^{(1)}_{n,i}] y^{(0)}_{n,i}$

Now the next step would be to find the first order correction to the compatible eigenvectors $y^{(1)}_{n,i}$. That is solve $[A_0 - \lambda^{(0)}_n] y^{(1)}_{n,i} = -[A_1 - \lambda^{(1)}_{n,i}] y^{(0)}_{n,i}$.
If you remember from last time, we used that since $\lambda^{(0)}_m \neq \lambda^{(0)}_n$ then we could smash it w/ $x^{(0)}_m$ (inner product), massage the result and obtain:

$$\langle x^{(0)}_{m,i}, y^{(1)}_{n,j} \rangle = \frac{\langle x^{(0)}_{m,i}, A_1 y^{(0)}_{n,j} \rangle}{\lambda^{(0)}_n - \lambda^{(0)}_m}$$

Then using this to determine the coefficients in: $y^{(1)}_{n,j} = \sum_i \langle x^{(0)}_{m,i}, y^{(1)}_{n,j} \rangle x^{(0)}_{m,i}$
we got our first order correction to the eigenvectors. $= \sum_i \frac{\langle x^{(0)}_{m,i}, A_1 y^{(0)}_{n,j} \rangle}{\lambda^{(0)}_n - \lambda^{(0)}_m} x^{(0)}_{m,i}$

Note $m=n$ is okay since $=0$ in this case as well!

But in the degenerate case we do have different vectors for which $m=n$, hence $\lambda^{(0)}_n - \lambda^{(0)}_n = 0$ ruins the story. But it turns out that just as using the existence of a solution to $[A_0 - \lambda^{(0)}_n] y^{(1)}_{n,i} = -[A_1 - \lambda^{(1)}_{n,i}] y^{(0)}_{n,i}$ ended up giving us info about $\lambda^{(1)}_n$ even before we knew $y^{(1)}_{n,i}$, then so to will demanding that: $[A_0 - \lambda^{(1)}_{n,i}] y^{(2)}_{n,i} = -[A_1 - \lambda^{(1)}_{n,i}] y^{(1)}_{n,i} + \lambda^{(2)}_{n,i} y^{(0)}_{n,i}$ have a solution will bring us info about $y^{(1)}_{n,i}$. Recall, demanding a solution only involves the r.h.s.!

So for $[A_0 - \lambda^{(1)}_{n,i}] y^{(2)}_{n,i} = -[A_1 - \lambda^{(1)}_{n,i}] y^{(1)}_{n,i} + \lambda^{(2)}_{n,i} y^{(0)}_{n,i}$ to have a solution we need the r.h.s. to be orthogonal to the subspace spanned by $x^{(0)}_{n,j}$.
$\langle x^{(0)}_{n,j}, [A_1 - \lambda^{(1)}_{n,i}] y^{(1)}_{n,i} - \lambda^{(2)}_{n,i} y^{(0)}_{n,i} \rangle = 0$ for all i and j
Using the expansion of $y^{(0)}_{n,i}$ in terms of $x^{(0)}_{n,j}$, i.e. $y^{(0)}_{n,i} = \sum_j a_{ij} x^{(0)}_{n,j}$
$a_{ij} \lambda^{(2)}_{n,i} = \langle x^{(0)}_{n,j}, [A_1 - \lambda^{(1)}_{n,i}] y^{(1)}_{n,i} \rangle$

Now we can split $y^{(1)}_{n,i}$ into two parts $y^{(1)}_{n,i} = y^{(1)\parallel}_{n,i} + y^{(1)\perp}_{n,i}$ where $y^{(1)\parallel}_{n,i}$ is the (unknown) part of $y^{(1)}_{n,i}$ which lives in the subspace spanned by $x^{(0)}_{n,j}$, and $y^{(1)\perp}_{n,i}$ is the rest (which we can get from the result near the top of this page.) and is orthogonal to the subspace.
Then we have:

$$\langle x^{(0)}_{n,j}, [A_1 - \lambda^{(1)}_{n,i}] y^{(1)\parallel}_{n,i} \rangle = -\underbrace{\langle x^{(0)}_{n,j}, A_1 y^{(1)\perp}_{n,i} \rangle}_{\text{since } \langle x^{(0)}_{n,j}, y^{(1)\perp}_{n,i} \rangle = 0} + a_{ij} \lambda^{(2)}_{n,i}$$

Now we can expand $y^{(1)\parallel}_{n,i}$ in terms of $x^{(0)}_{n,j}$ w/ $y^{(1)\parallel}_{n,i} = \sum_k b_{ik} x^{(0)}_{n,k}$

Putting this in we get:

$$\sum_k \left(x^{(0)}_{n,i,j},\ [A_1 - \lambda^{(1)}_{n,i}]\, x^{(0)}_{n,k}\right) b_{ik} = \lambda^{(2)}_{n,i}\, a_{ij} - \left(x^{(0)}_{n,j},\ A_1\, y^{(0)\perp}_{n,i}\right) \quad \text{for all } i \text{ and } j$$

Again this can be cast in a simpler matrix form: $(M - \lambda^{(1)}_{n,i})\, b_i = \lambda^{(2)}_{n,i}\, a_i - c_i$

$M_{jk} = (x^{(0)}_{n,i,j},\ A_1\, x^{(0)}_{n,k})$

$c_i = (x^{0}_{n,i,j},\ A_1\, y^{(1)\perp}_{n,i})$

This is again of the type $(M - \lambda)\, x = y$ against which we can smash it and get stuff out.

Example: $A_0 = \begin{pmatrix}1&0&0\\0&2&1\\0&1&2\end{pmatrix} \Rightarrow \underbrace{\lambda_1^{(0)} = 1,\ \lambda_2^{(0)} = 1}_{\lambda_{2,i}^{(0)}:\ i=1,2},\ \lambda_3^{(0)} = 3$ w/ $\underbrace{x_1^{(0)} = \begin{pmatrix}a\\b\\-b\end{pmatrix},\ x_2^{(0)} = \begin{pmatrix}c\\d\\-d\end{pmatrix}}_{\text{now we can choose what to start w/}},\ x_3^{(0)} = \begin{pmatrix}0\\1/\sqrt{2}\\1/\sqrt{2}\end{pmatrix}$

$$\underbrace{x_1^{(0)} = \begin{pmatrix}1/\sqrt{2}\\1/2\\-1/2\end{pmatrix}}_{x_{2,1}^{(0)}},\ \underbrace{x_2^{(0)} = \begin{pmatrix}1/\sqrt{2}\\-1/2\\1/2\end{pmatrix}}_{x_{2,2}^{(0)}} \Rightarrow (x_i^{(0)}, x_j^{(0)}) = \delta_{ij}$$

Now consider the perturbation by ϵA_1 w/ $A_1 = \begin{pmatrix}0&0&0\\0&0&1\\0&1&0\end{pmatrix}$.

First note that since $\lambda_3^{(0)}$ has a distinct eigenvector, then we can use:

$$\lambda_3^{(1)} = (x_3^{(0)}, A_1 x_3^{(0)}) = 1$$

$$x_3^{(1)} = \sum_{n\neq 3} \frac{(x_n^{(0)}, A_1 x_3^{(0)})}{\lambda_3^{(0)} - \lambda_n^{(0)}} x_n^{(0)} = \frac{(x_1^{(0)}, A_1 x_3^{(0)})}{\lambda_3^{(0)} - \lambda_1^{(0)}}\begin{pmatrix}1/\sqrt{2}\\1/2\\-1/2\end{pmatrix} + \frac{(x_2^{(0)}, A_1 x_3^{(0)})}{\lambda_3^{(0)} - \lambda_2^{(0)}}\begin{pmatrix}1/\sqrt{2}\\-1/2\\1/2\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix} \quad \text{since } A_1 x_3^{(0)} = x_3^{(0)}$$

$$\lambda_3^{(2)} = (x_3^{(0)}, A_1 x_3^{(1)}) = 0$$

For $\lambda_{2,i}^{(0)}$ consider $M_{jk} = \langle x_{2,j}^{(0)}, A_1 x_{2,k}^{(0)}\rangle \Rightarrow M = \begin{pmatrix}(x_{2,1}^{(0)}, A_1 x_{2,1}^{(0)}) & (x_{2,1}^{(0)}, A_1 x_{2,2}^{(0)})\\ (x_{2,2}^{(0)}, A_1 x_{2,1}^{(0)}) & (x_{2,2}^{(0)}, A_1 x_{2,2}^{(0)})\end{pmatrix} = \begin{pmatrix}-\frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & -\frac{1}{2}\end{pmatrix}$

We want to solve $(M - \lambda_{2,i}^{(1)}) a_i = 0 \Rightarrow \det(M - \lambda_{2,i}^{(1)}) = 0 = (-\frac{1}{2} - \lambda_{2,i}^{(1)})^2 - \frac{1}{4}$

$$0 = \lambda_{2,i}^{(1)} + \lambda_{2,i}^{(1)\,2} \Rightarrow \underbrace{\lambda_{2,1}^{(1)} = 0,\ \lambda_{2,2}^{(1)} = -1}_{\text{first order corrections to deg. eigenvalues}}$$

$\lambda_{2,1}^{(1)}$: $\begin{pmatrix}-\frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & -\frac{1}{2}\end{pmatrix}\begin{pmatrix}a\\b\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix} \Rightarrow \begin{matrix}-\frac{1}{2}a + \frac{1}{2}b = 0\\ \frac{1}{2}a - \frac{1}{2}b = 0\end{matrix} \Rightarrow \begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\end{pmatrix} = a_{1k}$

$\lambda_{2,2}^{(1)}$: $\begin{pmatrix}-\frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & -\frac{1}{2}\end{pmatrix}\begin{pmatrix}a\\b\end{pmatrix} = \begin{pmatrix}-a\\-b\end{pmatrix} \Rightarrow \begin{matrix}-\frac{1}{2}a + \frac{1}{2}b = -a\\ \frac{1}{2}a - \frac{1}{2}b = -b\end{matrix} \Rightarrow \begin{pmatrix}1/\sqrt{2}\\-1/\sqrt{2}\end{pmatrix} = a_{2k}$

So now we know that $y_{2,1}^{(0)} = \sum_k a_{1k} x_{2,k}^{(0)} = \frac{1}{\sqrt{2}}\begin{pmatrix}1/\sqrt{2}\\1/2\\-1/2\end{pmatrix} + \frac{1}{\sqrt{2}}\begin{pmatrix}1/\sqrt{2}\\-1/2\\1/2\end{pmatrix} = \begin{pmatrix}1\\0\\0\end{pmatrix}$

and

$$y_{2,2}^{(0)} = \sum_k a_{2k} x_{2,k}^{(0)} = \frac{1}{\sqrt{2}}\begin{pmatrix}1/\sqrt{2}\\1/2\\-1/2\end{pmatrix} - \frac{1}{\sqrt{2}}\begin{pmatrix}1/\sqrt{2}\\-1/2\\1/2\end{pmatrix} = \begin{pmatrix}0\\1/\sqrt{2}\\-1/\sqrt{2}\end{pmatrix}$$

Note: $\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1/\sqrt{2}\\-1/\sqrt{2}\end{pmatrix}, \begin{pmatrix}0\\1/\sqrt{2}\\1/\sqrt{2}\end{pmatrix}$ is also an orthonormal set! If we had guessed this right to begin with, it would have cleaned up the story.

Let's check: $A = \begin{pmatrix}1&0&0\\0&2&1+\epsilon\\0&1+\epsilon&2\end{pmatrix} \Rightarrow \det\begin{pmatrix}1-\lambda&0&0\\0&2-\lambda&1+\epsilon\\0&1+\epsilon&2-\lambda\end{pmatrix} = (1-\lambda)[(2-\lambda)^2 - (1+\epsilon)^2] \Rightarrow \lambda = 1$

$$\lambda^2 - 4\lambda + 4 - 1 - 2\epsilon - \epsilon^2$$

$$\lambda^2 - 4\lambda + (3 - 2\epsilon - \epsilon^2)$$

$$\lambda = \frac{4 \pm \sqrt{16 - 12 + 8\epsilon + 4\epsilon^2}}{2}$$

$$= 2 \pm \sqrt{\epsilon^2 + 2\epsilon + 1} = 2 \pm \sqrt{(\epsilon+1)^2}$$

$$\lambda = 2 \pm (\epsilon + 1) = 3 + \epsilon,\ 1 - \epsilon$$

So in all we have:

$\lambda = 1 \quad = \lambda_{2,1}$

$\lambda = 1 - \epsilon \quad = \lambda_{2,2}$

$\lambda = 3 + \epsilon \quad = \lambda_3$