Easy Perturbation Theory

Let $M(\varepsilon)$ be a one parameter family of matrices that depends smoothly on the parameter ε. That is, each matrix element of $M(\varepsilon)$ is a C^∞ function of ε. Suppose that it is known that the unit vector $\mathbf{v}_0$ is an eigenvector of $M(0)$ of eigenvalue λ_0 and that it is also known that the eigenvalue λ_0 is simple. Then $\det\big(M(0)-\lambda\mathbb{1}\big)$ has a simple zero at $\lambda=\lambda_0$. Since zeroes of polynomials depend continuously on the coefficients of the polynomials, there is a neighbourhood $\mathcal{O}$ of λ_0 such that $\det\big(M(\varepsilon)-\lambda\mathbb{1}\big)$ has exactly one zero in $\mathcal{O}$ for each sufficiently small ε. Hence, for each sufficiently small ε, $M(\varepsilon)$ has exactly one eigenvalue, $\lambda(\varepsilon)$, in $\mathcal{O}$. Of course the corresponding eigenvector, $\mathbf{v}(\varepsilon)$, is only determined up to a multiplicative constant, but we can select a unique eigenvector by requiring that, for example, the dot product $\big(\mathbf{v}(\varepsilon),\mathbf{v}_0\big)=1$. Thus

$$\big[M(\varepsilon)-\lambda(\varepsilon)\mathbb{1}\big]\mathbf{v}(\varepsilon)=\mathbf{0}\qquad\big(\mathbf{v}(\varepsilon),\mathbf{v}_0\big)=1$$

Differentiating with respect to ε gives

$$\big[M'(\varepsilon)-\lambda'(\varepsilon)\mathbb{1}\big]\mathbf{v}(\varepsilon)+\big[M(\varepsilon)-\lambda(\varepsilon)\mathbb{1}\big]\mathbf{v}'(\varepsilon)=\mathbf{0}\qquad\big(\mathbf{v}'(\varepsilon),\mathbf{v}_0\big)=0 \tag{1_ε}$$

or

$$M'(\varepsilon)\mathbf{v}(\varepsilon)+M(\varepsilon)\mathbf{v}'(\varepsilon)=\lambda'(\varepsilon)\mathbf{v}(\varepsilon)+\lambda(\varepsilon)\mathbf{v}'(\varepsilon)$$

Taking the inner product with $\mathbf{v}_0$, and exchanging the left and right hand sides of the equation,

$$\lambda'(\varepsilon)=\big(M'(\varepsilon)\mathbf{v}(\varepsilon)+M(\varepsilon)\mathbf{v}'(\varepsilon),\mathbf{v}_0\big) \tag{2_ε}$$

To simplify the coming computations, let's also assume that the matrix $M(0)$ is self–adjoint. That is $(M(0)\mathbf{v},\mathbf{w})=(\mathbf{v},M(0)\mathbf{w})$ for all vectors $\mathbf{v}$ and $\mathbf{w}$. Then, since $M(0)-\lambda_0\mathbb{1}$ maps the line $L=\{\alpha\mathbf{v}_0\}$ to zero (which is in the line), $M(0)-\lambda_0\mathbb{1}$ maps the orthogonal complement of the line, $L^\perp=\{\ \mathbf{v}\ \big|\ \mathbf{v}\perp\mathbf{v}_0\ \}$, to itself. Since λ_0 is a simple eigenvalue of $M(0)$, the dimension of the kernel of $M(0)-\lambda_0\mathbb{1}$ is exactly one and so the restriction of $M(0)-\lambda_0\mathbb{1}$ to $L^\perp$ must be one–to–one and hence invertible. Let, with abuse of notation, $\big[M(0)-\lambda_0\mathbb{1}\big]^{-1}$ denote the matrix whose restriction to L is zero and whose restriction to $L^\perp$ is the inverse of the restriction to $L^\perp$ of $M(0)-\lambda_0\mathbb{1}$. That is, $\big[M(0)-\lambda_0\mathbb{1}\big]^{-1}\mathbf{v}_0=\mathbf{0}$ and if $\mathbf{v}\perp\mathbf{v}_0$, then $\big[M(0)-\lambda_0\mathbb{1}\big]^{-1}\mathbf{v}$ is the unique $\mathbf{w}\in L^\perp$ obeying $\big[M(0)-\lambda_0\mathbb{1}\big]\mathbf{w}=\mathbf{v}$. The matrix $\big[M(0)-\lambda_0\mathbb{1}\big]^{-1}$ may be constructed as follows. Let $M(0)-\lambda_0\mathbb{1}=UDU^{-1}$ be a diagonalization of $M(0)-\lambda_0\mathbb{1}$. Thus D is a diagonal matrix with diagonal entries being the eigenvalues of $M(0)-\lambda_0\mathbb{1}$. Exactly one of these diagonal entries is zero. Let D' be the diagonal matrix with each diagonal entry being the inverse of the corresponding diagonal entry of D, except that the zero diagonal entry of D is left as is. Then $\big[M(0)-\lambda_0\mathbb{1}\big]^{-1} = UD'U^{-1}$.

Now, setting $\varepsilon = 0$ in (2_ε) and using

$$\big(M(0)\mathbf{v}'(0),\mathbf{v}_0\big) = \big(\mathbf{v}'(0),M(0)\mathbf{v}_0\big) = \big(\mathbf{v}'(0),\lambda_0\mathbf{v}_0\big) = \lambda_0\big(\mathbf{v}'(0),\mathbf{v}_0\big) = 0$$

gives

$$\lambda'(0) = \big(M'(0)\mathbf{v}_0,\mathbf{v}_0\big) \tag{2_0}$$

Setting $\varepsilon = 0$ in (1_ε) and subbing back in (2_0) gives

$$\big[M(0)-\lambda(0)\mathbb{1}\big]\mathbf{v}'(0) = -M'(0)\mathbf{v}_0 + \lambda'(0)\mathbf{v}_0 = -M'(0)\mathbf{v}_0 + \big(M'(0)\mathbf{v}_0,\mathbf{v}_0\big)\mathbf{v}_0$$

The right hand side is exactly the projection of $-M'(0)\mathbf{v}_0$ on $L^\perp$ and, in particular, is in $L^\perp$. Since $\big(\mathbf{v}'(0),\mathbf{v}_0\big) = 0$, $\mathbf{v}'(0)$ is itself in $L^\perp$ and

$$\mathbf{v}'(0) = -\big[M(0)-\lambda(0)\mathbb{1}\big]^{-1}\big[M'(0)\mathbf{v}_0+\big(M'(0)\mathbf{v}_0,\mathbf{v}_0\big)\mathbf{v}_0\big] = -\big[M(0)-\lambda(0)\mathbb{1}\big]^{-1}M'(0)\mathbf{v}_0 \tag{1_0}$$

Recall that, by definition, $\big[M(0)-\lambda(0)\mathbb{1}\big]^{-1}\mathbf{v}_0 = 0$. We now know $\lambda'(0)$ and $\mathbf{v}'(0)$.

Differentiating (2_ε) with respect to ε gives

$$\lambda''(\varepsilon) = \big(M''(\varepsilon)\mathbf{v}(\varepsilon) + 2M'(\varepsilon)\mathbf{v}'(\varepsilon) + M(\varepsilon)\mathbf{v}''(\varepsilon),\mathbf{v}_0\big)$$

Setting $\varepsilon = 0$ and using

$$\big(M(0)\mathbf{v}''(0),\mathbf{v}_0\big) = \big(\mathbf{v}''(0),M(0)\mathbf{v}_0\big) = \big(\mathbf{v}''(0),\lambda_0\mathbf{v}_0\big) = \lambda_0\big(\mathbf{v}''(0),\mathbf{v}_0\big) = 0$$

(the derivative of (1_ε) includes $\big(\mathbf{v}''(\varepsilon),\mathbf{v}_0\big) = 0$) and (1_0) gives

$$\begin{aligned}\lambda''(0) &= \big(M''(0)\mathbf{v}_0 + 2M'(0)\mathbf{v}'(0),\mathbf{v}_0\big)\\ &= \big(M''(0)\mathbf{v}_0,\mathbf{v}_0\big) - 2\big(M'(0)\big[M(0)-\lambda(0)\mathbb{1}\big]^{-1}M'(0)\mathbf{v}_0,\mathbf{v}_0\big)\end{aligned}$$

Continuing in this way, one can compute all derivatives, $\lambda^{(n)}(0)$ and $\mathbf{v}^{(n)}(0)$, of $\lambda(\varepsilon)$ amd $\mathbf{v}(\varepsilon)$ at $\varepsilon = 0$. If $M(\varepsilon)$ is analytic in ε at $\varepsilon = 0$, the same is true for $\lambda(\varepsilon)$ and $\mathbf{v}(\varepsilon)$ and the computed derivatives determine them.