Regularly perturbed matrix eigenvalue problems

Stilianos Louca

December 2012

What this is

This quick-and-dirty article provides an introduction to regularly perturbed matrix eigenvalue problems, to leading order approximation. This is an old topic with lots of literature available. A good introduction into the subject can also be found in [1, §1.6].

Problem statement

Let $\mathbb{K}$ be either $\mathbb{R}$ or $\mathbb{C}$. Let $A \in \mathbb{K}^{N\times N}$ be a square matrix with right-eigenvalue $\lambda_o \in \mathbb{K}$, of algebraic and geometric multiplicity n. Denote by $E_{\lambda_o}(A)$ the corresponding eigenspace. Let $x_1, .., x_n \in \mathbb{K}^N$ be a basis for $E_{\lambda_o}(A)$. Let $B \in \mathbb{K}^{N\times N}$ be another square matrix and let ε be a small parameter. We wish to find a leading order asymptotic solution for the right-eigenvalue problem

$$Ay = \lambda y + \varepsilon By, \quad \lambda \in \mathbb{K},\ y \in \mathbb{K}^N \tag{1}$$

of the form $y = y_o + \varepsilon y_1 + o(\varepsilon)$ (for some $y_o \in E_{\lambda_o}(A)$) and $\lambda = \lambda_o + \varepsilon\lambda_1 + o(\varepsilon)$, as $\varepsilon \to 0$.

Power series expansion

Inserting the expansions into (1) yields

$$Ay_o + \varepsilon Ay_1 + o(\varepsilon) = \lambda_o y_o + \varepsilon(\lambda_o y_1 + \lambda_1 y_o + By_o) + o(\varepsilon). \tag{2}$$

Comparing orders of ε leads to the conditions

$$Ay_o = \lambda_o y_o \tag{3}$$

and

$$(A - \lambda_o)y_1 = (\lambda_1 + B)y_o. \tag{4}$$

Condition (3) translates to

$$y_o = \sum_{i=1}^{n} \alpha_i x_i \tag{5}$$

for suitable coefficients $\alpha_1, .., \alpha_n \in \mathbb{K}$. Now let $\overline{x}_1, .., \overline{x}_n \in \mathbb{K}^N$ be a basis for kernel$(A^T - \lambda_o)$ (note that the latter has the same dimension as kernel$(A - \lambda_o)$). Then condition (4) implies

$$0 = \underbrace{\overline{x}_k^T (A - \lambda_o)}_{0} y_1 \overset{(4)}{=} \overline{x}_k^T \cdot (\lambda_1 + B)\, y_o \tag{6}$$

for all $k \in \{1, .., n\}$. Using (5) we can write (6) in the matrix form

$$\boxed{\lambda_1 Q\boldsymbol{\alpha} + R\boldsymbol{\alpha} = 0,} \tag{7}$$

where

$$\boldsymbol{\alpha} := \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}, \quad Q := \begin{pmatrix} \overline{x}_1^T \cdot x_1 & \dots & \overline{x}_1^T \cdot x_n \\ \vdots & \ddots & \vdots \\ \overline{x}_n^T \cdot x_1 & \dots & \overline{x}_n^T \cdot x_n \end{pmatrix}, \quad R := \begin{pmatrix} \overline{x}_1^T \cdot Bx_1 & \dots & \overline{x}_1^T \cdot Bx_n \\ \vdots & \ddots & \vdots \\ \overline{x}_n^T \cdot Bx_1 & \dots & \overline{x}_n^T \cdot Bx_n \end{pmatrix}. \tag{8}$$

Note that the *secular equation* (7) can be seen as a generalized eigenvalue problem for λ_1 and the coefficient vector $\boldsymbol{\alpha} \in \mathbb{K}^n$. Solving it yields the correct limit value $y_o = \lim_{\varepsilon\to 0} y$ (see (5)) and the correct first order correction factor λ_1. Subsequently solving (4) yields the correction factor y_1. Note that since $(A - \lambda_o)$ is non-injective (having a kernel of dimension n), the solutions y_1 will form an n-dimensional affine space in $\mathbb{K}^N$.

Sufficiency of conditions

Left to show is that the solutions $(\lambda_1, \boldsymbol{\alpha}) \in \mathbb{K} \times \mathbb{K}^N$ of (7) really allow for the original condition (4) to be satisfied through a suitable choice of y_1. Otherwise said, we wish to show that $(\lambda_1 + B)y_o$ is in the image of $(A - \lambda_o)$.

Reminder:

By elementary dual system theory we know that $\text{image}(A - \lambda_o) = \left[\text{kernel}(A^T - \lambda_o)\right]_\perp$, where we denote $U_\perp := \left\{x \in \mathbb{K}^N : z^T \cdot x = 0 \ \forall z \in U\right\}$ for any subspace U of $\mathbb{K}^N$. We denote $x \perp U$ for any $x \in U_\perp$.

It thus suffices to show that $(\lambda_1 + B)y_o \perp \text{kernel}(A^T - \lambda)$. For that it actually suffices to show $\overline{x}_k^T \cdot (\lambda_1 + B)y_o = 0$ for all $k \in \{1, \dots, n\}$, since $\overline{x}_1, \dots, \overline{x}_n$ is a basis for $\text{kernel}(A^T - \lambda)$. But this is trivially satisfied by (6), which is equivalent to (7).

Example: Derivative of eigenvalues

The above result can be readily applied to calculate the derivative(s) of an eigenvalue λ with respect to a parameter $\mathbf{q} \in \mathbb{R}^d$ when the matrix depends on $\mathbf{q}$. More precisely, let $A = A(\mathbf{q}) \in \mathbb{K}^{N\times N}$ depend sufficiently smoothly on $\mathbf{q}$, and denote by $B_i := \partial_i A$ the partial derivative of A with respect to q^i. By the above results, particularly Eq. (7), we know that $\partial_i\lambda$ must satisfy

$$(\partial_i\lambda) \cdot Q\boldsymbol{\alpha} - R_i\boldsymbol{\alpha} = 0 \tag{9}$$

for a suitable $\boldsymbol{\alpha} \in \mathbb{K}^n \setminus \{0\}$, where

$$\boldsymbol{\alpha} := \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}, \quad Q := \begin{pmatrix} \overline{x}_1^T \cdot x_1 & \dots & \overline{x}_1^T \cdot x_n \\ \vdots & \ddots & \vdots \\ \overline{x}_n^T \cdot x_1 & \dots & \overline{x}_n^T \cdot x_n \end{pmatrix}, \quad R_i := \begin{pmatrix} \overline{x}_1^T \cdot B_i x_1 & \dots & \overline{x}_1^T \cdot B_i x_n \\ \vdots & \ddots & \vdots \\ \overline{x}_n^T \cdot B_i x_1 & \dots & \overline{x}_n^T \cdot B_i x_n \end{pmatrix}. \tag{10}$$

Note that the generalized eigenvalue equation (9) might have multiple solutions, so that $\partial_i\lambda$ can have multiple values. In the special case $n = 1$ (i.e. λ is not degenerate) at the point at which the derivative is evaluated, Eq. (9) simplifies to

$$\boxed{\partial_i\lambda = \frac{\overline{x}^{\mathrm{T}}(\partial_i A)x}{\overline{x}^{\mathrm{T}}x},} \tag{11}$$

where $\overline{x}$ and x are left- and right-eigenvectors of A for the eigenvalue λ.

Alternative

Instead of deploying dual spaces, one could make use of the euclidean scalar product existing on $\mathbb{K}^n$. One would then instead choose $\overline{x}_1, \dots, \overline{x}_n$ to be a basis in $\text{kernel}(A^\dagger - \lambda_o^*)$, where we denote $A^\dagger := (A^T)^*$ (transposed & complex conjugate). Also, one would then replace all row-column products by the scalar product. The matrices Q, R defined in (8) would then take the form

$$Q := \begin{pmatrix} \langle\overline{x}_1, x_1\rangle & \dots & \langle\overline{x}_1, x_n\rangle \\ \vdots & \ddots & \vdots \\ \langle\overline{x}_n, x_1\rangle & \dots & \langle\overline{x}_n, x_n\rangle \end{pmatrix}, \quad R := \begin{pmatrix} \langle\overline{x}_1, Bx_1\rangle & \dots & \langle\overline{x}_1, Bx_n\rangle \\ \vdots & \ddots & \vdots \\ \langle\overline{x}_n, Bx_1\rangle & \dots & \langle\overline{x}_n, Bx_n\rangle \end{pmatrix}, \tag{12}$$

with condition (7) staying the same.

Special case for hermitian matrices

Suppose $A^\dagger = A$, then $\lambda_o \in \mathbb{R}$ and $(A^\dagger - \lambda_o^*) = (A - \lambda_o)$. One can choose $x_1, .., x_n$ to be orthonormal and set $\overline{x}_k = x_k \; \forall k$. The matrix Q defined in (8) would then simplify to the identity $Q = \mathrm{Id}$, so that the secular equation (7) simplifies to a simple eigenvalue equation

$$\boxed{\lambda_1 \boldsymbol{\alpha} + R\boldsymbol{\alpha} = 0,} \tag{13}$$

where

$$R := \begin{pmatrix} \langle x_1, Bx_1 \rangle & \dots & \langle x_1, Bx_n \rangle \\ \vdots & \ddots & \vdots \\ \langle x_n, Bx_1 \rangle & \dots & \langle x_n, Bx_n \rangle \end{pmatrix}, \tag{14}$$

Now suppose that B was also hermitian. Then R becomes hermitian as well, since $\langle x_k, Bx_j \rangle = \langle Bx_k, x_j \rangle = \langle x_j, Bx_k \rangle^*$. It is therefore diagonalizable, i.e. admitting n linearly independent eigenvectors $\boldsymbol{\alpha}^{(1)}, \dots, \boldsymbol{\alpha}^{(n)} \in \mathbb{C}^n$ with corresponding eigenvalues $\lambda_1^{(1)}, .., \lambda_1^{(n)} \in \mathbb{R}$. If some of those eigenvalues are unequal, then the perturbation εB results in a reduction of the degeneracy of the eigenspace $E_{\lambda_o}(A + \varepsilon B)$, resulting in multiple eigenvalue branches from λ_o.

References

[1] *Hinch, E.J. (1991)*, Perturbation Methods
Cambridge University Press