Perturbation method

Liping Liu

April 20, 2016

1 Algebraic problems

We are interested in finding the approximate roots of a polynomial of form:

$$P(x,\varepsilon)=\sum_{k=0}^{n}a_k^{(\varepsilon)}x^k=0 \qquad (a_n^{\varepsilon}\neq 0), \tag{1}$$

where the coefficients $a_k^{(\varepsilon)}$ is a polynomial of ε and $P(x,\varepsilon)$ is a polynomial of x of degree n if $\varepsilon\neq 0$. From the fundamental theorem of algebra, the above polynomial admits n-roots. Subsequently a root of the polynomial (1) is denoted by $r(\varepsilon)$. We are interested in the asymptotic behaviors of roots $r(\varepsilon)$ as $\varepsilon\to 0$.

Theorem 1 (Simmonds and Mann 1998) *Each root of (1), $r=r(\varepsilon)$, is continuous as a function of ε in a punctuated neighborhood at $\varepsilon=0$ and can be written as*

$$r(\varepsilon)=\varepsilon^p w(\varepsilon),$$

where p is called a proper value, $w(0)\neq 0$ and $\varepsilon\mapsto w(\varepsilon)$ is analytic in a neighborhood of $\varepsilon=0$:

$$w(\varepsilon)=\sum_{k=0}^{\infty}c_k\varepsilon^k. \tag{2}$$

To find the p-value, there are three different scenarios.

1. If the root $r(0)$ is a simple root (of multiplicity 1), we may set $p=0$ and insert (2) into (1). Then all coefficients c_k in (2) can be recursively determined. For example, consider

$$P(x,\varepsilon)=x^4-\varepsilon x-1=0.$$

Inserting (2) into the above equation, we have

$$\Big(\sum_{k=0}^{\infty}c_k\varepsilon^k\Big)^4-\varepsilon\sum_{k=0}^{\infty}c_k\varepsilon^k-1=0,$$

which implies that

$$\begin{aligned}
\varepsilon^0: &\qquad c_0^4-1=0, &&\Rightarrow c_0=1,i,-1,-i;\\
\varepsilon^1: &\qquad 4c_0^3c_1-c_0=0, &&\Rightarrow c_1=1/(4c_0^2);\\
\varepsilon^0: &\qquad 4c_0^3c_2+6c_0^2c_1-c_1=0, &&\Rightarrow c_2=(c_1-6c_0^2c_1)/(4c_0^3);\\
&\quad \ldots\ldots
\end{aligned}$$

2. Roots that blow up as $\varepsilon \to 0$. The proper value $p < 0$. We can figure out the p-value by dominant balance analysis. For example, consider

$$P(x, \varepsilon) = \varepsilon^2 x^6 - \varepsilon x^4 - x^3 + 8 = 0. \tag{3}$$

The four terms in the above equation scale as ε^q, $q = 6p + 2, 4p + 1, 3p, 0$, respectively. To achieve $P(x, \varepsilon) = 0$ as $\varepsilon \to 0$, *at least two out of these four terms need to "balance" each other whereas all other terms are negligibly small.* The possibilities are listed as follows:

   (a) If $6p+2 = 4p+1$, then $p = -1/2$ and the four terms in (3) scale as ε^q, $q = -1, -1, -3/2, 0$. The largest term is however the third term that scales as $\varepsilon^{-3/2}$, meaning a contradiction;

   (b)

(see detailed analysis on pg. 325, textbook of Bender and Orszag)

3. A repeated (or degenerate) root of multiplicity $k \geq 2$. The proper value $p > 0$ and we can similarly figure out the p-value by dominant balance analysis. For example, consider

$$P^{(\varepsilon)}(x) = x^2 - 2\varepsilon x - \varepsilon = 0.$$

(see detailed analysis on pg. 8, Example 5.3. in the note of D. Gilliam)

We remark that the first scenario is referred to as **regular perturbation** whereas the second and third scenarios are referred to as **singular perturbations**.

2 Transcendental equation

A transcendental equation is an equation containing transcendental functions, i.e., analytic functions that are not polynomials. For example, $x = e^x$ and $\tan x = x$. Closed-form solutions to such equations are rare. Nevertheless, the asymptotic behavior of transcendental equation can still be obtained near an exact solution under some mild conditions. One of the well-known situations is summarized as *Lagrange's inversion formula.*

Theorem 2 *Let $g(z) \neq 0$ be an analytic function in a neighborhood of $z = 0$. Then*

$$w = \frac{z}{g(z)}$$

implies that in a neighborhood of $w = 0$,

$$z = z(w) = \sum_{k=1}^{\infty} c_k w^k, \qquad c_k = \frac{1}{k!}\left(\frac{d}{dz}\right)^{k-1} (g(z))^k \bigg|_{z=0}.$$

The above theorem follows from a more general theorem called *implicit function theorem* or *open mapping theorem.* To see its applications, we consider the following examples.

1. Consider the equation $xe^x = t^{-1}$. Find the asymptotic behavior of $x = x(t)$ as $t \to +\infty$.

   (see detailed analysis on pg. 19, Example 5.7. in the note of D. Gilliam)

2. Consider the equation $x^t = e^{-x}$. Find the asymptotic behavior of $x = x(t)$ as $t \to +\infty$.

   (see detailed analysis on pg. 19, Example 5.8. in the note of D. Gilliam)

3. Consider the equation $\tan x = x$. Find the asymptotic behavior of the nth solution $x_n = x(n)$ as $n \to +\infty$.

   (see detailed analysis on pg. 20, Example 5.9. in the note of D. Gilliam)

3 Eigenvalue problems

We begin our discuss for linear operators defined on a finite dimensional space, i.e., symmetric real matrices or Hermitian matrices, and then discuss about Hermitian operators that may be regarded as matrices of infinite dimensions. First of all, we recall a theorem in linear algebra for real symmetric matrices and complex Hermitian matrices.

Theorem 3 *(i) For every real symmetric matrix $\mathbf{A} \in \mathbb{R}^{n\times n}_{\text{sym}}$, there exist n eigenvectors $\{\mathbf{e}_1, \cdots \mathbf{e}_n\} \in \mathbb{R}^n$ that form an orthonormal basis of $\mathbb{R}^n$. The analogous theorem for complex matrix is as follows: (ii) For every Hermitian matrix $\mathbf{A} \in \mathbb{C}^{n\times n}_{\text{Her}}$, there exist n eigenvectors eigenvectors $\{\mathbf{e}_1, \cdots \mathbf{e}_n\} \in \mathbb{C}^n$ that form an orthonormal basis of $\mathbb{C}^n$. In addition, all eigenvalues are real.*

3.1 Perturbation method for finite dimensional eigenvalue-eigenvector problems

Now let us consider a symmetric matrix $\mathbf{A}_0 \in \mathbb{R}^{n\times n}_{\text{sym}}$. By the above theorem, we know there exist an orthonormal basis $\{\mathbf{e}^0_1, \cdots \mathbf{e}^0_n\} \in \mathbb{R}^n$ formed by eigenvectors of $\mathbf{A}_0$:

$$\mathbf{A}_0\mathbf{e}^0_i = \lambda^0_i \mathbf{e}^0_i.$$

We are interested in finding the eigenvalues and eigenvectors of a new symmetric matrix $\mathbf{A} = \mathbf{A}_0 + \varepsilon\mathbf{A}_1$ that is a small perturbation of the original matrix $\mathbf{A}_0$. To this end, we can assume that the original eigenvalue-eigenvector pair $(\lambda^0_i, \mathbf{e}^0_i)$ becomes $(\lambda_i, \mathbf{e}_i)$ that can be expanded in term s of ε as:

$$(\lambda_i, \mathbf{e}_i) = (\lambda^0_i, \mathbf{e}^0_i) + \varepsilon(\lambda^1_i, \mathbf{e}^1_i) + \varepsilon^2(\lambda^2_i, \mathbf{e}^2_i) + \cdots .$$

Without loss of generality we can impose the normalization condition: $\mathbf{e}_i \cdot \mathbf{e}^0_i = 1$, i.e.,

$$\mathbf{e}^k_i \cdot \mathbf{e}^0_i = 0 \qquad \forall\, k \geq 1 \ \&\ i = 1, \cdots, n. \tag{4}$$

By definition we shall have

$$\mathbf{A}\mathbf{e}_i = (\mathbf{A}_0 + \varepsilon\mathbf{A}_1)(\mathbf{e}^0_i + \varepsilon\mathbf{e}^1_i + \varepsilon^2\mathbf{e}^2_i + \cdots) = (\lambda^0_i + \varepsilon\lambda^1_i + \varepsilon^2\lambda^2_i + \cdots)(\mathbf{e}^0_i + \varepsilon\mathbf{e}^1_i + \varepsilon^2\mathbf{e}^2_i + \cdots).$$

Arranging terms according to the order of ε, we obtain

$$\begin{aligned}
\varepsilon^0: &\qquad \mathbf{A}_0\mathbf{e}^0_i = \lambda^0_i\mathbf{e}^0_i, \\
\varepsilon^1: &\qquad \mathbf{A}_0\mathbf{e}^1_i + \mathbf{A}_1\mathbf{e}^0_i = \lambda^0_i\mathbf{e}^1_i + \lambda^1_i\mathbf{e}^0_i, \\
\varepsilon^2: &\qquad \mathbf{A}_0\mathbf{e}^2_i + \mathbf{A}_1\mathbf{e}^1_i = \lambda^0_i\mathbf{e}^2_i + \lambda^1_i\mathbf{e}^1_i + \lambda^2_i\mathbf{e}^0_i, \\
\vdots &\qquad \cdots\cdots
\end{aligned} \tag{5}$$

The first of the above equation is clearly trivial. Taking the inner product of the second equation with $\mathbf{e}^0_j$ we have

$$\lambda^0_j\mathbf{e}^0_j \cdot \mathbf{e}^1_i + \mathbf{e}^0_j \cdot \mathbf{A}_1\mathbf{e}^0_i = \lambda^0_i\mathbf{e}^0_j \cdot \mathbf{e}^1_i + \lambda^1_i\delta_{ij}. \tag{6}$$

The implication of the above equation has two separate cases. For clarity and without loss of generality, below we focus on the first eigenvalue-eigenvector, setting $i = 1$.

- Nondegenerate case. If $\lambda_j^0 \neq \lambda_1^0$ for $j \neq 1$, the above equation implies

$$\begin{cases} \lambda_1^1 = \mathbf{e}_1^0 \cdot \mathbf{A}_1 \mathbf{e}_1^0 & \text{if } j = 1; \\ \mathbf{e}_1^1 \cdot \mathbf{e}_j^0 = \frac{\mathbf{e}_j^0 \cdot \mathbf{A}_1 \mathbf{e}_1^0}{\lambda_1^0 - \lambda_j^0} & \text{if } j \neq 1. \end{cases}$$

Together with (4), the last of the above equation implies

$$\mathbf{e}_1^1 = \sum_{j \neq 1} \frac{\mathbf{e}_j^0 \cdot \mathbf{A}_1 \mathbf{e}_1^0}{\lambda_1^0 - \lambda_j^0} \mathbf{e}_j^0. \tag{7}$$

Moreover, taking the inner product of the third equation of (5) with $\mathbf{e}_j^0$ we have

$$\lambda_j^0 \mathbf{e}_j^0 \cdot \mathbf{e}_i^2 + \mathbf{e}_j^0 \cdot \mathbf{A}_1 \mathbf{e}_i^1 = \lambda_i^0 \mathbf{e}_j^0 \cdot \mathbf{e}_i^2 + \lambda_i^1 \mathbf{e}_j^0 \cdot \mathbf{e}_i^1 + \lambda_i^2 \delta_{ij}. \tag{8}$$

For the nondegenerate state $i = 1$, from the above equation we obtain

$$\begin{cases} \lambda_1^2 = \mathbf{e}_1^0 \cdot \mathbf{A}_1 \mathbf{e}_1^1 = \sum_{j \neq 1} \frac{(\mathbf{e}_j^0 \cdot \mathbf{A}_1 \mathbf{e}_1^0)^2}{\lambda_1^0 - \lambda_j^0} & \text{if } j = 1; \\ \mathbf{e}_1^2 \cdot \mathbf{e}_j^0 = \frac{\mathbf{e}_j^0 \cdot \mathbf{A}_1 \mathbf{e}_1^1 - \lambda_1^1 \mathbf{e}_j^0 \cdot \mathbf{e}_1^1}{\lambda_1^0 - \lambda_j^0} & \text{if } j \neq 1. \end{cases}$$

- Degenerate case. If $\lambda_j^0 \neq \lambda_1^0$ for $j > m$ and $\lambda_j^0 = \lambda_1$ for $j = 1, \cdots, m$, then the eigenvectors can be chosen to be any orthonormal basis of the eigenspace $V := \{\mathbf{e} : \mathbf{A}_0 \mathbf{e} = \lambda_1^0 \mathbf{e}\} = \text{span}\{\mathbf{e}_1^0, \cdots, \mathbf{e}_m^0\}$. Further, equation (6) implies that

$$\begin{cases} \lambda_1^1 \delta_{1j} = \mathbf{e}_j^0 \cdot \mathbf{A}_1 \mathbf{e}_1^0 & \text{if } j = 1, \cdots, m; \\ \mathbf{e}_1^1 \cdot \mathbf{e}_j^0 = \frac{\mathbf{e}_j^0 \cdot \mathbf{A}_1 \mathbf{e}_1^0}{\lambda_1^0 - \lambda_j^0} & \text{if } j > m. \end{cases} \tag{9}$$

The first of the above equation of course cannot be true for general eigenvectors in the eigenspace V. Instead, we can show that $(9)_1$ requires that $\mathbf{e}_1^0, \cdots, \mathbf{e}_m^0$ are orthonormal eigenvectors of the linear mapping $\mathbf{A}_1$ restricted to the subspace V, which is defined as

$$\begin{cases} \mathbf{A}_1' : & V \to V, \\ & \mathbf{A}_1' \mathbf{x} = \mathcal{P} \mathbf{A}_1 \mathbf{x}, \end{cases}$$

where $\mathcal{P} : \mathbb{R}^n \to V$ is the projection operator. In addition, it can be shown that the matrix representation of $\mathbf{A}_1'$ for an orthonormal basis $\{\mathbf{f}_1^0, \cdots, \mathbf{f}_m^0\}$ of V is given by

$$(\mathbf{A}_1')_{ij} = \mathbf{f}_j^0 \cdot \mathbf{A}_1 \mathbf{f}_i^0 \qquad i, j = 1, \cdots, m.$$

Then the eigenvectors $\mathbf{e}_1^0, \cdots, \mathbf{e}_m^0$ such that $(9)_1$ holds can be find by solving the eigenvalue problem:

$$\mathbf{A}_1' \mathbf{e}_i^0 = \lambda_i^1 \mathbf{e}_i^0 \qquad i = 1, \cdots, m,$$

which determines perturbed eigenvalues. In regard of $(9)_2$, the leading correction of eigenvector $\mathbf{e}_1^1$ is given by

$$\mathbf{e}_1^1 = \sum_{j > m} \frac{\mathbf{e}_j^0 \cdot \mathbf{A}_1 \mathbf{e}_1^0}{\lambda_1^0 - \lambda_j^0} \mathbf{e}_j^0.$$

3.2 *Perturbation method for infinite dimensional eigenvalue-eigenfunction problems

We now consider a linear operator $H : \text{Dom}(H) \to \overline{\mathcal{D}}$, where $\text{Dom}(H)$ is a *function space* (typically a dense subspace of $\overline{\mathcal{D}}$). Heuristically, we may regard linear operators (on an infinite dimensional vector space) as a matrix of infinite size. To apply the above theorems and calculations to operators, we need to properly generalize the inner product for vector space $\overline{\mathcal{D}}$ and the concept of "symmetric" operators.

To be concrete, for periodic functions $E(x), \rho(x)$ with period p we consider the Hermitian operator $H : \text{Dom}(H) \to \overline{\mathcal{D}}$ defined by

$$H(x) = -\frac{1}{\rho(x)}\frac{d}{dx}E(x)\frac{d}{dx},$$
$$\overline{\mathcal{D}} = \{f \in L^2_{loc}(\mathbb{R}) : f(x+p) = e^{ikp}f(x)\}.$$

- Inner product:

$$\langle f, g\rangle = \int_0^p \rho \bar{f} g dx.$$

- The operator H is Hermitian since

$$\langle f, Hg\rangle = \langle g, Hf\rangle \qquad \forall\ f, g \in \mathcal{D}.$$

- The operator H is positive since

$$\langle f, Hf\rangle > 0 \qquad \forall\ 0 \neq f \in \mathcal{D}.$$

Then in analogy with, we have the following theorem:

Theorem 4 *There exist $(\lambda_n, \phi_n) \in \mathbb{R} \times \overline{\mathcal{D}}$ $(n = 1, 2, \cdots)$ such that (i) $H\phi_n = \lambda_n\phi_n$ (i.e., eigenvalues and eigenfunctions), (ii) $\langle \phi_i, \phi_j\rangle = \delta_{ij}$, and (iii) for any function $f \in \overline{\mathcal{D}}$, we have*

$$\sum_{i=1}^{\infty}\langle f, \phi_n\rangle\phi_n \to f \qquad \text{in } \bar{D}.$$

We now consider a Hermitian operator $H_0 : \text{Dom}(H_0) \to \overline{\mathcal{D}}$. By the above theorem, we know there exist an orthonormal basis $\{\phi_1^0, \phi_2^0, \cdots\} \subset \overline{\mathcal{D}}$ formed by eigenfunctions of H_0:

$$H_0\phi_i^0 = \lambda_i^0\phi_i^0.$$

We are interested in finding the eigenvalues and eigenvectors of a new Hermitian operator

$$H = H_0 + \varepsilon H_1,$$

which is a small perturbation of the original operator H_0. To this end, we can assume that the original eigenvalue-eigenfunction pair (λ_i^0, ϕ_i^0) becomes (λ_i, ϕ_i) that can be expanded in term s of ε as:

$$(\lambda_i, \phi_i) = (\lambda_i^0, \phi_i^0) + \varepsilon(\lambda_i^1, \phi_i^1) + \varepsilon^2(\lambda_i^2, \phi_i^2) + \cdots .$$

Without loss of generality we can impose the normalization condition: $\langle \phi_i, \phi_i^0 \rangle = 1$, i.e.,

$$\langle \phi_i^k, \phi_i^0 \rangle = 0 \qquad \forall\, k \geq 1. \tag{10}$$

By definition we shall have

$$H\phi_i = (H_0 + \varepsilon H_1)(\phi_i^0 + \varepsilon\phi_i^1 + \varepsilon^2\phi_i^2 + \cdots) = (\lambda_i^0 + \varepsilon\lambda_i^1 + \varepsilon^2\lambda_i^2 + \cdots)(\phi_i^0 + \varepsilon\phi_i^1 + \varepsilon^2\phi_i^2 + \cdots).$$

Arranging terms according to the order of ε, we obtain

$$\begin{aligned} \varepsilon^0 : \qquad & H_0\phi_i^0 = \lambda_i^0\phi_i^0, \\ \varepsilon^1 : \qquad & H_0\phi_i^1 + H_1\phi_i^0 = \lambda_i^0\phi_i^1 + \lambda_i^1\phi_i^0, \\ \varepsilon^2 : \qquad & \\ & \cdots\cdots \end{aligned} \tag{11}$$

The first of the above equation is clearly trivial. Taking the inner product of the second equation with ϕ_j^0 we have

$$\lambda_j^0\langle \phi_j^0, \phi_i^1 \rangle + \langle \phi_j^0, H_1\phi_i^0 \rangle = \lambda_i^0\langle \phi_j^0, \phi_i^1 \rangle + \lambda_i^1\delta_{ij}. \tag{12}$$

The implication of the above equation again has two separate cases. For clarity and without loss of generality, below we focus on the first eigenvalue-eigenfunction, setting $i = 1$.

- Nondegenerate case. If $\lambda_j^0 \neq \lambda_1^0$ for $j \neq 1$, the above equation implies

$$\begin{cases} \lambda_1^1 = \langle \phi_1^0, \mathbf{A}_1\phi_1^0 \rangle & \text{if } j = 1; \\ \langle \phi_1^1, \phi_j^0 \rangle = \frac{\langle \phi_j^0, H_1\phi_1^0 \rangle}{\lambda_1^0 - \lambda_j^0} & \text{if } j \neq 1. \end{cases}$$

  Together with (10), the last of the above equation implies

$$\phi_1^1 = \sum_{j\neq 1} \frac{\langle \phi_j^0, H_1\phi_1^0 \rangle}{\lambda_1^0 - \lambda_j^0}\phi_j^0.$$

- Degenerate case. If $\lambda_j^0 \neq \lambda_1^0$ for $j > m$ and $\lambda_j^0 = \lambda_1$ for $j = 1, \cdots, m$, then the eigenfunctions of H_0 associated with eigenvalue λ_1^0 can be chosen to be any orthonormal basis of the eigenspace $\mathcal{V} := \{\phi : H_0\phi = \lambda_1^0\phi\} = \text{span}\{\phi_1^0, \cdots, \phi_m^0\}$. Further, equation (6) implies that

$$\begin{cases} \lambda_1^1\delta_{1j} = \langle \phi_j^0, H_1\phi_1^0 \rangle & \text{if } j = 1, \cdots, m; \\ \langle \phi_1^1, \phi_j^0 \rangle = \frac{\langle \phi_j^0, H_1\phi_1^0 \rangle}{\lambda_1^0 - \lambda_j^0} & \text{if } j > m. \end{cases} \tag{13}$$

  The first of the above equation of course cannot be true for general eigenfunctions in the eigenspace $\mathcal{V}$. Instead, we can show that $(13)_1$ requires that $\mathbf{e}_1^0, \cdots, \mathbf{e}_m^0$ are orthonormal eigenvectors of the linear mapping H_1 restricted to the subspace $\mathcal{V}$, which is defined as

$$\begin{cases} H_1' : & \mathcal{V} \to \mathcal{V}, \\ & H_1'\phi = \mathcal{P}H_1\phi, \end{cases}$$

where $\mathcal{P} : \bar{D} \to \mathcal{V}$ is the projection operator. In addition, it can be shown that the matrix representation of H_1' for an orthonormal basis $\{\varphi_1^0, \cdots, \varphi_m^0\}$ of $\mathcal{V}$ is given by

$$(H_1')_{ij} = \langle \varphi_j^0, H_1 \varphi_i^0 \rangle \qquad i, j = 1, \cdots, m.$$

Then the eigenfunctions $\phi_1^0, \cdots, \phi_m^0$ such that $(13)_1$ holds can be find by solving the eigenvalue problem:

$$H_1' \phi_i^0 = \lambda_i^1 \phi_i^0 \qquad i = 1, \cdots, m,$$

which determines perturbed eigenvalues. In regard of $(13)_2$, the leading correction of eigenfunction ϕ_1^1 is given by

$$\phi_1^1 = \sum_{j>m} \frac{\langle \phi_j^0, H_1 \phi_1^0 \rangle}{\lambda_1^0 - \lambda_j^0} \phi_j^0.$$

4 Integral evaluations

Typical techniques of evaluating an integral that depends on a small parameter ε include:

1. Expansion of integrands;
2. Integration by parts;
3. Laplace method;
4. Method of stationary phase.

The most important examples that are relevant to physics/statistical mechanics concern about the *Gaussian integrals.*

4.1 Expansion of integrands

$$I(\varepsilon) = \int_0^1 \sin(\varepsilon x^2) dx \quad (\varepsilon \ll 1); \qquad \int_0^x t^{-3/4} e^{-t} dt \quad (x \ll 1).$$

4.2 Integration by parts

$$I(x) = \int_x^\infty \frac{e^{-t}}{t^2} dt \quad (x \gg 1); \qquad \int_0^\infty e^{-st} f(t) dt \quad (s \gg 1).$$

4.3 Laplace method

Assume $h = h(t)$ is a smooth function on the interval $[a, b]$ and achieves its maximum at $x = c$. Then

$$\begin{aligned} I(x) &= \int_a^b e^{xh(t)} f(t) dt \qquad (x \gg 1) \\ &\approx I(x, \varepsilon) = \begin{cases} \int_a^{a+\varepsilon} e^{-xh(t)} f(t) dt & \text{if } c = a, \\ \int_{c-\varepsilon}^{c+\varepsilon} e^{-xh(t)} f(t) dt & \text{if } c \in (a, b), \\ \int_{b-\varepsilon}^{b} e^{-xh(t)} f(t) dt & \text{if } c = b, \end{cases} \\ &= \begin{cases} -\frac{f(a) e^{xh(a)}}{xh'(a)} & \text{if } c = a \ \&\ h'(a) \neq 0, \\ \frac{f(c) e^{xh(c)} \sqrt{2\pi}}{[-xh''(c)]^{1/2}} & \text{if } c \in (a, b) \ \&\ h''(c) < 0, \\ \frac{f(b) e^{xh(b)}}{xh'(b)} & \text{if } c = b \ \&\ h'(b) \neq 0. \end{cases} \end{aligned} \tag{14}$$

Remark 5 *In statistical physics, we often encounter integral of such form*

$$F(\beta) = \frac{1}{Z} \int_a^b e^{-\beta H(t)} f(t) dt \qquad (\beta = 1/k_B T \gg 1),$$

where H is the Hamiltonian of the system, $Z = \int_a^b e^{-\beta H(t)} dt$ is the partition function, and the above integral yield the expected value of the thermodynamic quantity $F(\beta)$ in the low temperature *limit.*

Examples ($x \gg 1$):

$$I(x) = \int_0^{10} \frac{e^{-xt}}{1+t} dt$$

Watson's Lemma: If $f(t) = t^\alpha \sum_{n=0}^\infty a_n t^{\beta n}$ ($\alpha > -1, \beta > 0$) and $|f(t)| < Me^{ct}$ for some $M, c > 0$. Then

$$I(x) = \int_0^b f(t) e^{-xt} dt \sim \sum_{n=0}^\infty \frac{a_n \Gamma(\alpha + \beta n + 1)}{x^{\alpha+\beta n+1}}.$$

4.4 Method of stationary phase

Consider the oscillatory integral for $g \in C_0^\infty(\mathbb{R})$ and $x \gg 1$:

$$I(x) = \int_{\mathbb{R}} f(t) e^{ixg(t)} dt$$

Assume that

$$g'(c) = 0, \qquad g''(c) \neq 0.$$

We have

$$\begin{aligned} I(x) &= e^{ixg(c)} \int_{\mathbb{R}} f(t) e^{ix[g(t)-g(c)]} dt \approx e^{ixg(c)} \int_{c-\varepsilon}^{c+\varepsilon} f(t) e^{\frac{g''(c)}{2} ix(t-c)^2} dt \\ &\approx e^{ixg(c)} f(c) \int_{\mathbb{R}} e^{\frac{g''(c)}{2} ix(t-c)^2} dt = e^{ixg(c)} f(c) \sqrt{\frac{2\pi i}{xg''(c)}}. \end{aligned}$$

5 Ordinary differential equations

5.1 Regular perturbation

Consider a general first-order linear ODE for $\mathbf{y} : \mathbb{R} \to \mathbb{R}^n$:

$$\mathbf{y}'(x) = \mathbf{F}(x)\mathbf{y}(x), \qquad \mathbf{y}(0) = \mathbf{y}_0, \tag{15}$$

where $\mathbf{F} : \mathbb{R} \to \mathbb{R}^{n\times n}$ is Lipschitz continuous in the sense that $\exists M > 0$,

$$|\mathbf{F}(x) - \mathbf{F}(y)| \leq M|x - y| \qquad \forall\, x, y \in \mathbb{R}.$$

The the solution to (15) is given by

5.2 Boundary layer method

If the highest-order term of an ODE vanishes as $\varepsilon \to 0$, we are in the regime of singular perturbation. Unlike regular perturbation, a naive Taylor expansion with respect to the small parameter cannot work since the unperturbed problem does not even admit a solution. To see what happens, it is useful to study two simple constant coefficient equations:

$$\begin{aligned} &\varepsilon y'' - y' = 0, \qquad y(0) = 0, \quad y(1) = 1; \\ &\varepsilon^2 y'' + (1+\varepsilon^2)y' + y = 0, \qquad y(0) = 0, \quad y(1) = 1. \end{aligned} \tag{16}$$

Formal procedure of boundary layer method:

- We first solve $(16)_1$.[1] For outer region $(0, 1-\delta)$, neglecting $\varepsilon y''$ we obtain

  $$y' = 0, \qquad y(0) = 0,$$

  which implies solution

  $$y = 0 \qquad \text{on } (0, 1-\delta).$$

  For the inner region $(1-\delta, 1)$, we anticipate $y' \sim 1/\varepsilon$, $y'' \sim 1/\varepsilon^2$. To extract the asymptotic behavior in the inner region, we have change of variables that zoom in the region:

  $$Y(X) = y(x), \qquad X - 1 = \frac{x-1}{\varepsilon}.$$

  The above definition of $Y(X)$ implies that

  $$\frac{d}{dx}y(x) = \frac{d}{dx}Y(X) = \frac{d}{\varepsilon dX}Y(X), \qquad \frac{d^2}{dx^2}y(x) = \frac{1}{\varepsilon^2}\frac{d^2}{dX^2}Y(X).$$

  Therefore, $(16)_1$ implies that

  $$Y_{XX} - Y_X = 0, \qquad Y(1) = 1,$$

  and hence

  $$Y(X) = C_1 + C_2 e^{X-1}, \qquad C_1 + C_2 = 1.$$

  We need a second condition to fix both C_1 and C_2. To this end, we assume $(x-1) = a\varepsilon^{1/2}$ for some constant $a < 0$. Then $X - 1 = a\varepsilon^{-1/2} \to -\infty$ as $\varepsilon \to 0$. As $\varepsilon \to 0$, on one hand the inner solution $y(1 + a\varepsilon^{1/2}) = Y(1 + a\varepsilon^{-1/2}) \to C_1$; on the other hand the outer solution implies $y(1 + a\varepsilon^{1/2}) \to 0$. To be self-consistent, we shall equate two limits:

  $$C_1 = 0.$$

  In summary, we obtain our boundary-layer solution:

  $$y(x) = \begin{cases} 0 & \text{on } (0, 1-\varepsilon), \\ e^{(x-1)/\varepsilon} & \text{on } (\varepsilon, 1). \end{cases}$$

---

[1] The exact solution is given by

$$y(x) = \frac{e^{x/\varepsilon} - 1}{e^{1/\varepsilon} - 1}.$$

For more general problems, exact solutions are rarely possible.

- We next solve $(16)_2$.[2] For outer region $(\delta, 1)$, we anticipate $y \sim y' \sim y'' \sim 1$. Neglecting $O(\varepsilon)$-terms we obtain

$$y' + y = 0, \qquad y(1) = 1,$$

which implies solution

$$y = e^{-x+1} \qquad \text{on } (\delta, 1).$$

For the inner region $(0, \delta)$, we anticipate $y' \sim 1/\varepsilon^2$, $y'' \sim 1/\varepsilon^4$. To extract the asymptotic behavior in the inner region, we have change of variables that zoom in the region:

$$Y(X) = y(x), \qquad X = \frac{x}{\varepsilon^2}.$$

The above definition of $Y(X)$ implies that

$$\frac{d}{dx}y(x) = \frac{d}{dx}Y(X) = \frac{d}{\varepsilon^2 dX}Y(X), \qquad \frac{d^2}{dx^2}y(x) = \frac{1}{\varepsilon^4}\frac{d^2}{dX^2}Y(X).$$

Therefore, $(16)_2$ implies that

$$Y_{XX} + Y_X = 0, \qquad Y(0) = 0,$$

and hence

$$Y(X) = C_1 + C_2 e^{-X}, \qquad C_1 + C_2 = 1.$$

We need a second condition to fix both C_1 and C_2. To this end, we assume $x = a\varepsilon^{1/2}$ for some constant $a > 0$. Then $X = a\varepsilon \to +\infty$ as $\varepsilon \to 0$. As $\varepsilon \to 0$, on one hand the inner solution $y(a\varepsilon) = Y(a\varepsilon^{-1}) \to C_1$; on the other hand the outer solution implies $y(a\varepsilon) \to e$. To be self-consistent, we shall equate two limits:

$$C_1 = e.$$

In summary, we obtain our boundary-layer solution:

$$y(x) = \begin{cases} e - e^{1-x/\varepsilon^2} & \text{on } (0, \delta), \\ e^{1-x} & \text{on } (\delta, 1), \end{cases}$$

where δ can be determined by [3]

$$e - e^{1-\delta/\varepsilon^2} = e^{1-\delta} \Rightarrow \delta \sim \varepsilon^2.$$

From the above formal solution procedures to (16), we may raise the questions: (i) Why boundary layer occurs? (ii) What is the *right* scaling of a boundary layer? and (iii) Where is the boundary layer? Below we heuristically address these questions which will be useful for solving more general non-constant coefficient problems.

[2] The exact solution is given by

$$y(x) = \frac{e^{-x/\varepsilon^2} - e^{-x}}{e^{-1/\varepsilon^2} - e^{-1}}.$$

[3] The precise value of δ is of no importance.

Why boundary layer occurs?

First of all, the phenomenon of a boundary layer is not the only possible behavior in a singular perturbation problem. It occur if (i) on average the general solution behave like $e^{\lambda(\varepsilon)x}$ for some $\lambda(\varepsilon) \in \mathbb{R}$, and (ii) $\lambda(\varepsilon) \to \pm\infty$ as $\varepsilon \to 0$.

If $\lambda(\varepsilon)$ has nonzero imaginary part, the solution will be fast oscillating.

What is the scaling of a boundary layer?

There are two ways to determine the scaling: (i) The scaling can be determined by solve the constant coefficient equation at one of the boundary point, or (ii) By conjecturing a scaling $X - a = (x-a)/\delta$, upon changing of variables we shall obtain a *nontrivial* equation where at least two terms in the original equation remain.

Where is the boundary layer?

Assume that boundary layer occurs near one boundary point. The singular solution is given by $e^{\lambda(\varepsilon)x}$. If $\lambda(\varepsilon) \to +\infty$ as $\varepsilon \to 0$, then to guarantee a bounded solution on the interval $(0,1)$, the boundary layer shall occur on the right boundary point as for $(16)_1$; if $\lambda(\varepsilon) \to -\infty$ as $\varepsilon \to 0$, then to guarantee a bounded solution on the interval $(0,1)$, the boundary layer shall occur on the left boundary point as for $(16)_2$.

Example 1 Find solutions to the following problems by the boundary layer method:

$$\varepsilon y'' + a(x)y' + b(x)y(x) = 0 \qquad \text{on } (0,1), \qquad y(0) = A,\ y(1) = B,$$

where $a(x) > 0$ on $(0,1)$.

Example 2 Find solutions to the following problems by the boundary layer method:

$$\varepsilon y'' + a(x)y' + b(x)y(x) = 0 \qquad \text{on } (0,1), \qquad y(0) = A,\ y(1) = B,$$

where $a(x) < 0$ on $(0,1)$.

5.3 WKB method

Consider

$$\varepsilon y'' + y = 0, \qquad y(0) = 0, y(1) = 1.$$

A quick analysis shows that the singular roots are given by $\lambda(\varepsilon) = \pm\varepsilon^{-1/2}i$, and therefore, we expect solutions of form $y(x) \sim c_1 \sin x/\sqrt{\varepsilon} + c_2 \cos x/\sqrt{\varepsilon}$. Indeed, the exact solution is given by

$$y(x) = \frac{\sin(x/\sqrt{\varepsilon})}{\sin(1/\sqrt{\varepsilon})}.$$

This kind of behavior of solutions cannot be captured by boundary layers since there are singular oscillations in the interior domain.

More generally, let us consider

$$\varepsilon^2 y'' = Q(x)y, \qquad Q(x) \neq 0.$$

If $Q(x) > 0$, we anticipate boundary layers; if $Q(x) < 0$, we anticipate singular oscillations in the interior domain.

To capture such singular behaviors, we may consider ansartz of form

$$y(x) \sim \exp(\frac{1}{\delta}\sum_{n=0}^{\infty}\delta^n S_n(x)).$$

Inserting the above ansartz into equation, we find that

$$\begin{aligned} y' &\sim \frac{1}{\delta}\sum_{n=0}^{\infty}\delta^n S_n' y, \\ y'' &\sim \{\frac{1}{\delta}\sum_{n=0}^{\infty}\delta^n S_n'' + [\frac{1}{\delta}\sum_{n=0}^{\infty}\delta^n S_n']^2\}y, \end{aligned}$$

and henceforth,

$$\frac{\varepsilon^2}{\delta^2}(S_0')^2 + \frac{2\varepsilon^2}{\delta}S_0'S_1' + \frac{\varepsilon^2}{\delta}S_0'' + \cdots = Q(x).$$

By dominant balance we necessarily have $\delta \sim \varepsilon$, and

$$\begin{aligned} (S_0')^2 &= Q(x), \\ 2S_0'S_1' + S_0'' &= 0, \\ 2S_0'S_n' + S_{n-1}'' + \sum S_j'S_{n-j}' &= 0, \end{aligned}$$

Excercise

1. For small ε, find the first two terms in the expansion of each of the roots of
$$P(x, \varepsilon) = \varepsilon x^4 - x^3 + 3x - 2 = 0.$$

2. Determine a two term expansion for the large roots of $x = \tan x$ and $x \tan x = 1$.

3. Determine a two term expansion for the each of roots of
$$P(x, \varepsilon) = x^3 - (3 + \varepsilon)x - 2 + \varepsilon = 0.$$
(Hint: the zeroth order solution $x^3 - 3x - 2 = 0$ is given by $x = -1, -1, 2$.)

4. (Optional) Find the first two terms in the expansion of the roots of $x^3 - \varepsilon x^2 - \varepsilon^2 = 0$.

5. Consider the matrix $\mathbf{A}_\varepsilon = \mathbf{A}_0 + \varepsilon \mathbf{A}_1$ for some small number $\varepsilon << 1$.
$$\mathbf{A}_0 = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \qquad \mathbf{A}_1 = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$
Find the first three terms in the expansion of each eigenvalue-eigenvector pairs of $\mathbf{A}_\varepsilon$.

6. Evaluate the Fourier transformation of a function $f \in C_0^\infty(\mathbb{R}^+)$:
$$I(\omega) = \int_0^\infty e^{i\omega t} f(t) dt \quad (\omega \gg 1).$$
(see Example 5.16 in the note of D. G.)

7. Evaluate the integral
$$I(x) = \int_0^5 \frac{e^{-xt}}{(1 + t^2)} \qquad (x \gg 1).$$
(see Example 5.19 in D.G.)

8. Read remark 5.4 in the note of D.G. and write the proof for (14).