



UNIVERSIDAD DE SEVILLA

Trabajo Fin de Máster

**Método POD y aplicaciones: caso
discreto y continuo**

Presentado por:

Manuel Bernardino del Pino

Dirigido por:

Macarena Gómez Mármol

Samuele Rubino

DEPARTAMENTO DE ECUACIONES DIFERENCIALES Y ANÁLISIS
NUMÉRICO

MARZO DE 2021

A mi abuelo...

Abstract

The proper orthogonal decomposition (POD) is a technique that aims at obtaining a compact representation of the data reducing the computational cost. The key idea of the POD is to reduce a large number of interdependent variables to a much smaller number of uncorrelated variables while retaining as much as possible of the variation in the original variables. This method has been gaining popularity over the years and it is being used to detect the number of signals in a multichannel time-series, for forecasting in meteorology or for the design of gurney flaps, among others. In this work, a theoretical study is carried out on this method in both the continuous and the discrete cases, the effective calculation is exposed and two numerical simulations are made; one for each of the two cases mentioned above.

Índice general

Introducción	3
1. Resultados preliminares	5
1.1. Optimización no lineal	5
1.2. Espacios de Hilbert. Teoría espectral	8
2. El método POD en el espacio \mathbb{R}^m	12
2.1. Planteamiento del problema	12
2.2. La descomposición en valores singulares (SVD)	13
2.3. Relación entre el método POD y la descomposición en valores singulares (SVD)	14
2.4. Propiedades de la base POD	21
2.5. El método POD para productos escalares con peso	27
2.6. Algunas aplicaciones del método POD	34
2.6.1. Ecuaciones diferenciales ordinarias	34
2.6.2. Ecuaciones en derivadas parciales	47
3. Aplicaciones prácticas del método POD	50
3.1. Caso discreto	50
3.2. Caso continuo	53
Bibliografía	59

Introducción

El método proper orthogonal decomposition (POD) es una técnica que tiene como objetivo obtener una representación compacta de una serie de datos reduciendo el coste computacional. La idea clave de este método es reducir un gran número de variables interdependientes a un número mucho menor de variables no correlacionadas mientras se retiene la “información esencial” de las variables originales. Se realiza una transformación ortogonal a la base de los autovectores de una cierta matriz, y los datos se proyectan en el subespacio generado por los autovectores correspondientes a los autovalores más grandes. Esencialmente, proporciona una base ortogonal para representar un conjunto dado de datos en un cierto sentido óptimo de mínimos cuadrados, es decir, ofrece formas de encontrar aproximaciones óptimas de dimensiones inferiores para ese conjunto de datos.

Este método, también conocido como la descomposición de Karhunen-Loève (KLD), fue propuesto de forma independiente por varios científicos, entre los que destacan Karhunen, Kosambi, Loève, Obukhov y Pougachev y fue concebido originalmente en el marco de procesos continuos de segundo orden. Cuando se restringe a un caso de dimensión finita y se trunca después de algunos términos, el POD es equivalente al análisis de componentes principales (ACP). Esta última metodología se originó con el trabajo de Pearson como un medio de ajustar planos por mínimos cuadrados ortogonales, pero también fue propuesta por Hotelling. Con esto quiero enfatizar que en la literatura el método POD aparece con distintos nombres dependiendo del área de aplicación, a saber, ACP en la literatura estadística, función ortogonal empírica en oceanografía y meteorología y análisis factorial en psicología y economía. En las referencias [8], [9] y [14] pueden encontrarse una discusión detallada sobre la equivalencia de POD, ACP y su conexión con la descomposición de valores singulares.

El método POD ha ido ganando popularidad con los años y se está utilizando en numerosos campos. Por ejemplo, Chacón et al. [3] lo usaron para hacer modelado computacional sobre gurney flaps y microtabs. Wax y Kailath [13] sugirieron su uso para detectar el número de señales en una serie de tiempo multicanal. Bayly et al. [2] empleó el POD para describir cambios cuantitativos en la complejidad espacial durante episodios prolongados de fibrilación ventricular. Epureanu et al. [4] obtuvieron modelos de orden reducido de flujos viscosos inestables. Barnston y Ropelewski [1] emplearon el método para obtener pronósticos en meteorología. Leen et al. [7] lo introdujeron como un medio para clasificar los datos de voz.

Este trabajo está formado por los siguientes capítulos:

En el primer capítulo presentamos algunos resultados sobre optimización no lineal y teoría espectral de operadores con el objeto de que esta memoria sea lo más autocontenida posible. Estos conceptos no son nuevos, ya que se han estudiado en diversas asignaturas del grado en matemáticas y en el máster.

En el segundo capítulo, que engloba la mayor parte del trabajo, se hace un desarrollo del método POD en los casos continuo y discreto. Veremos su relación con la descomposición en valores singulares, cómo se realiza el cálculo efectivo para encontrar una base POD, cuáles son algunas propiedades de estas bases y lo aplicaremos a algunas situaciones de interés. En concreto a una ecuación diferencial ordinaria y a una ecuación en derivadas parciales.

En el tercer, y último capítulo, vamos a realizar dos simulaciones numéricas. Una correspondiente al caso discreto y otra al caso continuo. En el caso discreto, el parámetro que se está usando para la POD es la temperatura interior. Concretamente, disponemos de datos experimentales correspondientes a las temperaturas en el interior de diversos patios y nuestro objetivo es usar esos valores para aproximar en otro patio nuevo. Con respecto al caso continuo, el parámetro que vamos a usar es el tiempo y vamos a aplicar el método POD en el caso concreto de un sistema diferencial ordinario de tres ecuaciones.

Capítulo 1

Resultados preliminares

En este capítulo vamos a recoger algunas herramientas fundamentales que usaremos a lo largo del trabajo sobre optimización no lineal y teoría espectral de operadores de modo que este sea lo más autocontenido posible. Para la elaboración del capítulo, hemos seguido fundamentalmente las referencias de [6] y [10].

1.1. Optimización no lineal

Vamos a plantearnos el problema de hallar los posibles máximos de una función, en principio no lineal, con restricciones de igualdad, es decir, queremos resolver:

$$(P) \quad \text{máx } J(x) \quad \text{sujeto a} \quad e(x) = 0,$$

donde $J : \mathbb{R}^n \rightarrow \mathbb{R}$ denota la *función objetivo* y $e : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $m \leq n$ son las *restricciones de igualdad*. Un punto $x \in \mathbb{R}^n$ se dice *admisibile* si $e(x) = 0$. El conjunto de las soluciones admisibles se define como

$$\mathcal{F}(P) = \{x \in \mathbb{R}^n : e(x) = 0\}.$$

Definición 1.1.1. Sea $\bar{x} \in \mathbb{R}^n$ un vector arbitrario.

- a) El punto \bar{x} se dice que es una *solución local* de (P) si $\bar{x} \in \mathcal{F}(P)$ y $J(\bar{x}) \geq J(x)$ para todo $x \in \mathcal{U}(\bar{x}) \cap \mathcal{F}(P)$, donde $\mathcal{U}(\bar{x}) \subset \mathbb{R}^n$ es un entorno abierto no vacío de \bar{x} .

- b) El punto \bar{x} se dice que es una *solución local estricta* de (P) si $\bar{x} \in \mathcal{F}(P)$ y $J(\bar{x}) > J(x)$ para todo $x \in \mathcal{U}(\bar{x}) \cap \mathcal{F}(P)$, donde $\mathcal{U}(\bar{x}) \subset \mathbb{R}^n$ es un entorno abierto no vacío de \bar{x} .
- c) El punto \bar{x} se dice que es una *solución global* de (P) si $\bar{x} \in \mathcal{F}(P)$ y $J(\bar{x}) \geq J(x)$ para todo $x \in \mathcal{F}(P)$.
- d) El punto \bar{x} se dice que es una *solución global estricta* de (P) si $\bar{x} \in \mathcal{F}(P)$ y $J(\bar{x}) > J(x)$ para todo $x \in \mathcal{F}(P)$.

Para caracterizar las soluciones de (P) necesitamos la noción de plano tangente. Una *curva* en un hiperplano $\mathcal{H} \subset \mathbb{R}^n$ es una familia de puntos $x(t) \in \mathcal{H}$, donde $x : [a, b] \rightarrow \mathcal{H}$ es continua y $a < b$. Diremos que la curva es *diferenciable* en t cuando $\dot{x}(t)$ exista. De igual forma, diremos que es *dos veces diferenciable* cuando exista $\ddot{x}(t)$. Diremos que la curva x *pasa por el punto* $\bar{x} \in \mathcal{H}$ si existe $\bar{t} \in [a, b]$ tal que $x(\bar{t}) = \bar{x}$. El conjunto de todos los vectores tangentes $\dot{x}(\bar{t})$ de todas las curvas diferenciales que pasan por \bar{x} se denomina *plano tangente* en \bar{x} .

Definición 1.1.2. Un punto $\bar{x} \in \mathcal{F}(P)$ se dice *regular* con respecto a la restricción $e(x) = 0$ si el conjunto $\{\nabla e_1(\bar{x}), \dots, \nabla e_m(\bar{x})\}$ es linealmente independiente en \mathbb{R}^n .

El siguiente resultado nos da una caracterización del plano tangente.

Teorema 1.1.3. Supongamos que $\bar{x} \in \mathcal{F}(P)$ es un punto regular. Entonces, el plano tangente en \bar{x} es igual al conjunto

$$\ker \nabla e(\bar{x}) = \{v \in \mathbb{R}^n : \nabla e(\bar{x})v = 0\} \subset \mathbb{R}^n,$$

donde

$$\nabla e(\bar{x}) = \begin{pmatrix} \nabla e_1(\bar{x})^\top \\ \vdots \\ \nabla e_m(\bar{x})^\top \end{pmatrix} \in \mathbb{R}^{m \times n}$$

es el Jacobiano de e en \bar{x} .

Ahora podemos formular las condiciones necesarias de optimalidad de primer orden para el problema (P).

Teorema 1.1.4 (Condiciones necesarias de optimalidad de primer orden). Supongamos que J y e son funciones continuamente diferenciables. Además, sea \bar{x} una solución local de (P) y un punto regular para $e(x) = 0$. Entonces, existe un único multiplicador de Lagrange $\bar{\lambda} = (\bar{\lambda}_1, \dots, \bar{\lambda}_m) \in \mathbb{R}^m$ resolviendo

$$\nabla J(\bar{x}) + \sum_{i=1}^m \bar{\lambda}_i \nabla e_i(\bar{x}) = \nabla J(\bar{x}) + \nabla e(\bar{x})^\top \bar{\lambda} = 0. \quad (1.1)$$

Nota 1.1.5. Si definimos la función Lagrangiana $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ como

$$\mathcal{L}(x, \lambda) = J(x) + \langle \lambda, e(x) \rangle_{\mathbb{R}^m} = J(x) + \lambda^\top e(x),$$

entonces (1.1) puede expresarse como

$$\nabla_x \mathcal{L}(\bar{x}, \bar{\lambda}) = \nabla J(\bar{x}) + \nabla e(\bar{x})^\top \bar{\lambda} = 0 \in \mathbb{R}^n.$$

Además, la restricción se satisface en \bar{x} , así que se cumple que

$$\nabla_\lambda \mathcal{L}(\bar{x}, \bar{\lambda}) = e(\bar{x}) = 0 \in \mathbb{R}^m.$$

Combinando estas dos igualdades, tenemos un sistema con $n + m$ ecuaciones para los vectores desconocidos $\bar{x} \in \mathbb{R}^n$ y $\bar{\lambda} \in \mathbb{R}^m$.

Si J y e son más regulares, podemos formular condiciones necesarias y suficientes de optimalidad de segundo orden.

Teorema 1.1.6 (Condiciones necesarias de optimalidad de segundo orden). Supongamos que J y e son dos veces diferenciable con continuidad. Además, sea \bar{x} una solución local de (P) y un punto regular para $e(x) = 0$. Entonces, la matriz $n \times n$

$$\nabla_{xx}^2 \mathcal{L}(\bar{x}, \bar{\lambda}) = \nabla^2 J(\bar{x}) + \sum_{i=1}^m \bar{\lambda}_i \nabla^2 e_i(\bar{x})$$

es semidefinida negativa en el conjunto $\ker \nabla e(\bar{x}) \subset \mathbb{R}^n$, es decir,

$$v^\top \nabla_{xx}^2 \mathcal{L}(\bar{x}, \bar{\lambda}) v \leq 0 \quad \text{para todo } v \in \ker \nabla e(\bar{x}).$$

Aquí, $\bar{\lambda} = (\bar{\lambda}_1, \dots, \bar{\lambda}_m)^\top \in \mathbb{R}^m$ denota al único multiplicador de Lagrange introducido en el Teorema 1.1.4.

Para asegurar que un punto $\bar{x} \in \mathcal{F}(P)$ es una solución de (P) tenemos que garantizar las condiciones suficientes de optimalidad.

Teorema 1.1.7 (Condiciones suficientes de optimalidad de segundo orden). Supongamos que J y e son dos veces diferenciable con continuidad. Además, sea el par $(\bar{x}, \bar{\lambda}) \in \mathbb{R}^n \times \mathbb{R}^m$ verificando las condiciones necesarias de optimalidad:

$$\nabla_x \mathcal{L}(\bar{x}, \bar{\lambda}) = 0,$$

$$\nabla_\lambda \mathcal{L}(\bar{x}, \bar{\lambda}) = 0,$$

Más aún, si \bar{x} es un punto regular para $e(x) = 0$, entonces la matriz $\nabla_{xx}^2 \mathcal{L}(\bar{x}, \bar{\lambda})$ es definida negativa en el conjunto $\ker \nabla e(\bar{x}) \subset \mathbb{R}^n$, es decir,

$$v^\top \nabla_{xx}^2 \mathcal{L}(\bar{x}, \bar{\lambda}) v < 0 \quad \text{para todo } v \in \ker \nabla e(\bar{x}) \setminus \{0\}.$$

1.2. Espacios de Hilbert. Teoría espectral

El objetivo de esta sección es recordar algunos conceptos de análisis funcional ligados a espacios de Hilbert así como una breve introducción a la teoría espectral de operadores.

Consideremos dos espacios de Hilbert $(\mathcal{H}_1, \langle \cdot, \cdot \rangle_{\mathcal{H}_1})$ y $(\mathcal{H}_2, \langle \cdot, \cdot \rangle_{\mathcal{H}_2})$. En lo que sigue, escribiremos simplemente \mathcal{H}_1 y \mathcal{H}_2 por abreviar la notación.

Definición 1.2.1. Sea $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$, es decir, una aplicación lineal y continua entre ambos espacios. Llamamos *imagen* o *rango* del operador T al conjunto

$$\text{Im } T = \{Tx : x \in \mathcal{H}_1\}.$$

Si $\text{Im } T$ resulta ser un subespacio de \mathcal{H}_2 de dimensión finita, diremos que el operador T es de *rango finito*.

Veamos que los operadores de rango finito admiten una representación muy sencilla.

Teorema 1.2.2. Supongamos que $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ tiene rango n . Entonces existen vectores v_1, \dots, v_n en \mathcal{H}_1 y vectores $\varphi_1, \dots, \varphi_n$ en \mathcal{H}_2 tales que para cada $x \in \mathcal{H}_1$

$$Tx = \sum_{j=1}^n \langle x, v_j \rangle \varphi_j.$$

Además, los vectores $\varphi_1, \dots, \varphi_n$ pueden ser tomados de forma que sean una base ortonormal de $\text{Im } T$.

Observemos que la representación de operadores de rango finito no es única pues dependiendo de la base ortonormal que escojamos tendremos distintas representaciones. Recordemos, a continuación, el concepto de operador adjunto en un espacio de Hilbert.

Definición 1.2.3. Sea $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$. Definimos el *operador adjunto* de T como la aplicación

$$\begin{aligned} T^* : \mathcal{H}_2 &\longrightarrow \mathcal{H}_1 \\ y &\longmapsto T^*y := y^* \end{aligned}$$

verificando $\langle Tx, y \rangle_{\mathcal{H}_2} = \langle x, T^*y \rangle_{\mathcal{H}_1}$, para cada $x \in \mathcal{H}_1$ y cada $y \in \mathcal{H}_2$.

Seguidamente describimos algunas propiedades notables de este operador.

Proposición 1.2.4. Sean $T, S \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ y $\alpha \in \mathbb{R}$. Entonces, se verifican las siguientes propiedades:

1. T^* es un operador lineal y continuo de \mathcal{H}_2 en \mathcal{H}_1 con $\|T^*\| = \|T\|$.
2. $(T + S)^* = T^* + S^*$
3. $(\alpha T)^* = \alpha T^*$.
4. $T^{**} = T$
5. Sea $D \in \mathcal{L}(\mathcal{H}_2, \mathcal{H}_3)$ entonces, $(DT)^* = T^*D^*$.

Definición 1.2.5. Sea $T \in \mathcal{L}(\mathcal{H})$. Diremos que T es *autoadjunto* si $T = T^*$.

Introducimos a continuación un subconjunto de los operadores continuos que necesitaremos para desarrollar el teorema espectral, los llamados operadores compactos.

Definición 1.2.6. Un operador $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ es *compacto* si para cada sucesión $\{x_n\}_{n=1}^{\infty}$ acotada en \mathcal{H}_1 , la sucesión $\{Tx_n\}_{n=1}^{\infty}$ tiene una subsucesión convergente en \mathcal{H}_2 .

Veamos algunos resultados importantes sobre este tipo de operadores:

Teorema 1.2.7. Un operador $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ es compacto si y solo si su operador adjunto $T^* \in \mathcal{L}(\mathcal{H}_2, \mathcal{H}_1)$ lo es.

Teorema 1.2.8. Sea $\{T_n\}$ una sucesión de operadores compactos en $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ y $T \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ tal que $\|T_n - T\| \xrightarrow{n \rightarrow \infty} 0$. Entonces T es un operador compacto.

Teorema 1.2.9. Sean $T : \mathcal{H} \rightarrow \mathcal{H}$ un operador compacto y $S \in \mathcal{L}(\mathcal{H})$. Entonces TS y ST son compactos.

Teorema 1.2.10. Sea $T \in \mathcal{L}(\mathcal{H})$ un operador de rango finito, entonces T es compacto.

Se vio en Algebra Lineal que si \mathcal{H} es de dimensión finita, entonces todo operador simétrico es diagonalizable, es decir, tiene una base de autovectores:

Teorema 1.2.11. Sea \mathcal{H} es un espacio euclídeo de dimensión n y $T \in \mathcal{L}(\mathcal{H})$ un operador autoadjunto, entonces existe una base ortonormal $\{\psi_1, \dots, \psi_n\}$ de \mathcal{H} y unos números reales $\lambda_1, \dots, \lambda_n$ tales que $T\psi_i = \lambda_i\psi_i, \quad \forall i = 1, \dots, n$.

El objetivo ahora será generalizar este resultado al caso de un operador autoadjunto sobre un espacio de Hilbert que puede ser infinito dimensional.

Definición 1.2.12. Sea $T \in \mathcal{L}(\mathcal{H})$. Diremos que un escalar $\lambda \in \mathbb{R}$ es un *autovalor* si existe $x \neq 0$ en \mathcal{H} tal que $Tx = \lambda x$. Al vector x se le llama *autovector* asociado al autovalor λ .

Es claro que $\lambda \in \mathbb{R}$ es un autovalor de T si y solo si $\text{Ker}(T - \lambda Id) \neq \{0\}$, es decir, si y solo si $T - \lambda Id$ no es inyectivo.

Los autovalores de un operador autoadjunto verifican las siguientes propiedades:

Proposición 1.2.13. Sea $T \in \mathcal{L}(\mathcal{H})$ un operador autoadjunto, entonces sus autovalores (en caso de tenerlos) son reales.

Proposición 1.2.14. Los autovalores asociados a distintos autovalores de un operador autoadjunto son ortogonales.

Vamos a terminar con un resultado clave en la teoría espectral: todo operador compacto y autoadjunto es diagonalizable.

Teorema 1.2.15. Sea $T \in \mathcal{L}(\mathcal{H})$ un operador compacto y autoadjunto. Entonces, existe un sistema ortonormal $\{x_n\}$ de autovectores de T y su correspondiente sucesión de autovalores $\{\lambda_n\}$ tal que para cada $x \in \mathcal{H}$ se tiene que

$$Tx = \sum_n \lambda_n \langle x, x_n \rangle x_n.$$

La sucesión $\{\lambda_n\}$ es decreciente y si es infinita converge a 0.

Capítulo 2

El método POD en el espacio \mathbb{R}^m

2.1. Planteamiento del problema

El método proper orthogonal decomposition, a partir de ahora, método POD, es una técnica que permite calcular una aproximación de unos vectores $y_1, \dots, y_n \in \mathbb{R}^m$ utilizando una serie de datos experimentales de los que se extrae su “información esencial” en términos de unos pocos vectores, que llamaremos base POD. Es decir, se trata de sustituir las variables originales por un número menor de variables que sean una combinación lineal de las originales. Supongamos que $y_1, \dots, y_n \in \mathbb{R}^m$ son los vectores dados que queremos aproximar, entonces el problema de aproximar simultáneamente estos vectores por $\ell (< n)$ vectores normalizados y ortogonales dos a dos, puede ser planteado como el siguiente problema de maximización:

$$\max_{\tilde{\psi}_1, \dots, \tilde{\psi}_\ell \in \mathbb{R}^m} \sum_{i=1}^{\ell} \sum_{j=1}^n |\langle y_j, \tilde{\psi}_i \rangle|^2 \quad \text{sujeto a} \quad \langle \tilde{\psi}_i, \tilde{\psi}_j \rangle = \delta_{ij} \text{ para } 1 \leq i, j \leq \ell,$$

donde δ_{ij} es la función delta de Kronecker, es decir, toma el valor uno si $i = j$ y cero en caso contrario.

El objetivo, por tanto, será resolver este problema de optimización no lineal, obtener algunas propiedades de la base POD y aplicar este método a alguna situación de interés. Para ello, la descomposición en valores singulares será una herramienta fundamental.

Para el desarrollo de este capítulo hemos seguido la referencia de [12].

2.2. La descomposición en valores singulares (SVD)

Sea $Y = (y_1 | \cdots | y_n) \in \mathbb{R}^{m \times n}$ una matriz de rango $d \leq \min\{m, n\}$ con las columnas $y_j \in \mathbb{R}^m$ para todo $j = 1, \dots, n$.

Recordemos que la descomposición en valores singulares, en adelante, SVD, de Y garantiza que existen unos números reales $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_d > 0$, una matriz ortogonal $\Psi \in \mathbb{R}^{m \times m}$ con columnas $\{\psi_i\}_{i=1}^m$ y $\Phi \in \mathbb{R}^{n \times n}$ ortogonal con columnas $\{\phi_i\}_{i=1}^n$ tales que:

$$\Psi^\top Y \Phi = \begin{pmatrix} D & \Theta \\ \Theta & \Theta \end{pmatrix} := \Sigma \in \mathbb{R}^{m \times n}, \quad (2.1)$$

donde $D = \text{diag}(\sigma_1, \dots, \sigma_d) \in \mathbb{R}^{d \times d}$, los ceros Θ denotan matrices nulas de dimensiones adecuadas y ‘ \top ’ la traspuesta de una matriz (o vector) (ver [5], pág 76).

Además, se puede probar fácilmente a partir de la definición (2.1) que los vectores $\{\phi_i\}_{i=1}^d$ y $\{\psi_i\}_{i=1}^d$ verifican que

$$Y \phi_i = \sigma_i \psi_i \quad \text{y} \quad Y^\top \psi_i = \sigma_i \phi_i \quad \text{para todo} \quad i = 1, \dots, d. \quad (2.2)$$

Además, utilizando estas igualdades obtenemos:

- $Y^\top Y \phi_i = \sigma_i Y^\top \psi_i = \sigma_i^2 \phi_i$ para todo $i = 1, \dots, d$.
- $Y Y^\top \psi_i = \sigma_i Y \phi_i = \sigma_i^2 \psi_i$ para todo $i = 1, \dots, d$.

Por tanto, $\{\phi_i\}_{i=1}^d$ y $\{\psi_i\}_{i=1}^d$ son el conjunto de autovectores de la matriz $Y^\top Y$ e $Y Y^\top$, respectivamente, asociado a los autovalores $\lambda_i = \sigma_i^2 > 0$. Los vectores $\{\phi_i\}_{i=d+1}^n$ (si $d < n$) y $\{\psi_i\}_{i=d+1}^m$ (si $d < m$) son autovectores de $Y^\top Y$ e $Y Y^\top$ asociados al autovalor nulo.

Por otra parte, de (2.1) y de las propiedades de Ψ y Φ , se obtiene que Y puede ser descompuesta como

$$Y = \Psi \Sigma \Phi^\top.$$

Es más, Y puede ser escrita como sigue:

$$Y = \Psi^d D (\Phi^d)^\top, \quad (2.3)$$

donde las matrices $\Psi^d \in \mathbb{R}^{m \times d}$ y $\Phi^d \in \mathbb{R}^{n \times d}$ vienen dadas por

$$\Psi_{ij}^d = \Psi_{ij} \text{ para } 1 \leq i \leq m; 1 \leq j \leq d,$$

$$\Phi_{ij}^d = \Phi_{ij} \text{ para } 1 \leq i \leq n; 1 \leq j \leq d.$$

Definiendo $B^d := D(\Phi^d)^\top \in \mathbb{R}^{d \times n}$, podemos reescribir (2.3) de la forma $Y = \Psi^d B^d$. Por tanto, el espacio columna de Y puede ser representado en términos de las d columnas independientes de Ψ^d . Los coeficientes de la expansión para las columnas y_j ($j = 1, \dots, n$) en la base $\{\psi_i\}_{i=1}^d$ vienen dados por la j -ésima columna de B^d . Además, como Ψ es ortogonal, se tiene:

$$\begin{aligned} y_j &= \sum_{i=1}^d B_{ij}^d \Psi_{\cdot,i}^d = \sum_{i=1}^d (D(\Phi^d)^\top)_{ij} \psi_i = \sum_{i=1}^d \underbrace{((\Psi^d)^\top \Psi^d)}_{=I_d} D(\Phi^d)^\top_{ij} \psi_i \\ &= \sum_{i=1}^d ((\Psi^d)^\top Y)_{ij} \psi_i = \sum_{i=1}^d \left(\underbrace{\sum_{k=1}^m \Psi_{ki}^d Y_{kj}}_{=\psi_i^\top y_j} \right) \psi_i = \sum_{i=1}^d \langle \psi_i, y_j \rangle_{\mathbb{R}^m} \psi_i, \end{aligned}$$

donde I_d es la matriz identidad de orden d , $\Psi_{\cdot,i}^d$ denota la i -ésima columna de la matriz Ψ^d y $\langle \cdot, \cdot \rangle_{\mathbb{R}^m}$ representa el producto escalar euclídeo en \mathbb{R}^m que denotamos a partir de ahora simplemente por $\langle \cdot, \cdot \rangle$ para simplificar la notación. Luego, hemos escrito los vectores columna de Y como combinación lineal de $\{\psi_i\}_i^d$:

$$y_j = \sum_{i=1}^d \langle y_j, \psi_i \rangle \psi_i \quad \text{para todo } j = 1, \dots, n.$$

2.3. Relación entre el método POD y la descomposición en valores singulares (SVD)

Supongamos que $y_1, \dots, y_n \in \mathbb{R}^m$ son los vectores dados que queremos aproximar. Para comenzar, vamos a considerar el problema de aproximar simultáneamente todos estos vectores y_j por uno sólo normalizado:

$$(\mathbf{P}^1) \quad \max_{\tilde{\psi} \in \mathbb{R}^m} \sum_{j=1}^n |\langle y_j, \tilde{\psi} \rangle|^2 \quad \text{sujeto a} \quad \|\tilde{\psi}\|^2 = 1,$$

donde $\|\tilde{\psi}\| = \sqrt{\langle \tilde{\psi}, \tilde{\psi} \rangle}$ para $\tilde{\psi} \in \mathbb{R}^m$.

Notemos que (\mathbf{P}^1) es un problema de optimización no lineal con restricción de igualdad que puede ser resuelto considerando las condiciones necesarias de optimalidad de primer orden. Para ello, vamos a escribir nuestro problema en la forma estándar definiendo la función $e : \mathbb{R}^m \rightarrow \mathbb{R}$ por $e(\psi) = 1 - \|\psi\|^2$ para $\psi \in \mathbb{R}^m$. Entonces, la restricción de (\mathbf{P}^1) puede ser expresada como $e(\psi) = 0$. Observemos que cualquier solución ψ de (\mathbf{P}^1) es no nula, ya que tiene que verificar $\|\psi\| = 1$. Por tanto, para cualquier solución ψ del problema, el conjunto $\{\nabla e(\psi)\} = \{-2\psi^\top\}$ es linealmente independiente y por ende cualquier solución de (\mathbf{P}^1) es un punto regular.

Sea $\mathcal{L} : \mathbb{R}^m \times \mathbb{R} \rightarrow \mathbb{R}$ el funcional Lagrangiano asociado a (\mathbf{P}^1) , es decir,

$$\mathcal{L}(\psi, \lambda) = \sum_{j=1}^n |\langle y_j, \psi \rangle|^2 + \lambda(1 - \|\psi\|^2) \quad \text{para } (\psi, \lambda) \in \mathbb{R}^m \times \mathbb{R}.$$

Supongamos que $\psi \in \mathbb{R}^m$ es una solución de (\mathbf{P}^1) . Como ψ es un punto regular, se deduce del Teorema 1.1.4 que existe un único multiplicador de Lagrange $\lambda \in \mathbb{R}$ verificando la condición necesaria de optimalidad de primer orden

$$\nabla \mathcal{L}(\psi, \lambda) = 0 \quad \text{en } \mathbb{R}^m \times \mathbb{R}.$$

Vamos a empezar calculando el gradiente de \mathcal{L} con respecto a ψ :

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \psi_i}(\psi, \lambda) &= \frac{\partial}{\partial \psi_i} \left(\sum_{j=1}^n \left| \sum_{k=1}^m Y_{kj} \psi_k \right|^2 + \lambda \left(1 - \sum_{k=1}^m \psi_k^2 \right) \right) \\ &= 2 \sum_{j=1}^n \left(\sum_{k=1}^m Y_{kj} \psi_k \right) Y_{ij} - 2\lambda \psi_i \\ &= 2 \sum_{k=1}^m \left(\underbrace{\sum_{j=1}^n Y_{ij} Y_{jk}^\top}_{=(YY^\top)_{ik}} \psi_k \right) - 2\lambda \psi_i. \end{aligned}$$

Por tanto, $\nabla_{\psi} \mathcal{L}(\psi, \lambda) = 2(YY^\top \psi - \lambda \psi) = 0$ si y solo si $YY^\top \psi = \lambda \psi$. Ahora bien, como la matriz YY^\top es simétrica y semidefinida positiva, existen m autovalores $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m \geq 0$ y autovectores asociados, que denotamos por $\{\psi_i\}_{i=1}^m$, que pueden ser elegidos de forma que el conjunto $\{\psi_1, \dots, \psi_m\}$ sea ortonormal. Es decir, se verifica

que $YY^\top \psi_i = \lambda_i \psi_i$ ($1 \leq i \leq m$) y además, si $\psi \in \mathbb{R}^m$ es un vector arbitrario, entonces

$$\psi = \sum_{i=1}^m \langle \psi, \psi_i \rangle \psi_i.$$

Por otra parte, la condición $\frac{\partial \mathcal{L}}{\partial \lambda}(\psi, \lambda) = 0$ nos conduce a $\|\psi\| = 1$. Es decir, la condición necesaria de optimalidad nos lleva a resolver el problema:

$$\begin{cases} YY^\top \psi = \lambda \psi & \text{en } \mathbb{R}^m, \\ \|\psi\| = 1 & \text{en } \mathbb{R}. \end{cases} \quad (2.4)$$

Debido a la SVD, en particular, el vector ψ_1 resuelve (2.4) y se tiene:

$$\begin{aligned} \sum_{j=1}^n |\langle y_j, \psi_1 \rangle|^2 &= \sum_{j=1}^n \langle y_j, \psi_1 \rangle \langle y_j, \psi_1 \rangle = \sum_{j=1}^n \langle \langle y_j, \psi_1 \rangle y_j, \psi_1 \rangle = \left\langle \sum_{j=1}^n \langle y_j, \psi_1 \rangle y_j, \psi_1 \right\rangle \\ &= \left\langle \sum_{j=1}^n \left(\sum_{k=1}^m Y_{kj}(\psi_1)_k \right) y_j, \psi_1 \right\rangle = \left\langle \sum_{k=1}^m \left(\sum_{j=1}^n Y_{kj} Y_{kj}^\top (\psi_1)_k \right), \psi_1 \right\rangle \\ &= \left\langle YY^\top \psi_1, \psi_1 \right\rangle = \lambda_1 \langle \psi_1, \psi_1 \rangle = \lambda_1 \underbrace{\|\psi_1\|^2}_{=1} = \lambda_1, \end{aligned}$$

donde $(\psi_1)_k$ es la k -ésima componente del vector ψ_1 e $Y_{.j}$ denota la j -ésima columna de la matriz Y . Observemos que $\nabla_{\psi\psi} \mathcal{L}(\psi, \lambda) = 2(YY^\top - \lambda I_m) \in \mathbb{R}^{m \times m}$.

Si evaluamos en (ψ_1, λ_1) se concluye de $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m \geq 0$ que

$$\begin{aligned} \langle \psi, \nabla_{\psi\psi} \mathcal{L}(\psi_1, \lambda_1) \psi \rangle &= 2 \left\langle \psi, (YY^\top - \lambda_1 I_m) \psi \right\rangle \\ &= 2 \sum_{i=1}^m \sum_{j=1}^m \langle \psi, \psi_i \rangle \langle \psi, \psi_j \rangle \left\langle \psi_1, (YY^\top - \lambda_1 I_m) \psi_j \right\rangle \\ &= 2 \sum_{i=1}^m \sum_{j=1}^m (\lambda_j - \lambda_1) \langle \psi, \psi_i \rangle \langle \psi, \psi_j \rangle \langle \psi_i, \psi_j \rangle \\ &= 2 \sum_{i=1}^m (\lambda_i - \lambda_1) |\langle \psi, \psi_i \rangle|^2 \leq 0. \end{aligned}$$

Por tanto, (ψ_1, λ_1) verifican las condiciones necesarias de optimalidad de segundo orden para un máximo, pero no las suficientes (Teoremas 1.1.6 y 1.1.7). Probemos a continuación que en realidad ψ_1 resuelve el problema (\mathbf{P}^1). En efecto, supongamos que

$\tilde{\psi} \in \mathbb{R}^m$ es un vector arbitrario con $\|\tilde{\psi}\| = 1$. Como $\{\psi_i\}_{i=1}^m$ es una base ortonormal de \mathbb{R}^m , tenemos:

$$\tilde{\psi} = \sum_{i=1}^m \langle \tilde{\psi}, \psi_i \rangle \psi_i.$$

Por tanto:

$$\begin{aligned} \sum_{j=1}^n |\langle y_j, \tilde{\psi} \rangle|^2 &= \sum_{j=1}^n \left| \left\langle y_j, \sum_{i=1}^m \langle \tilde{\psi}, \psi_i \rangle \psi_i \right\rangle \right|^2 \\ &= \sum_{j=1}^n \sum_{i=1}^m \sum_{k=1}^m \left(\langle y_j, \langle \tilde{\psi}, \psi_i \rangle \psi_i \rangle \langle y_j, \langle \tilde{\psi}, \psi_k \rangle \psi_k \rangle \right) \\ &= \sum_{j=1}^n \sum_{i=1}^m \sum_{k=1}^m \left(\langle y_j, \psi_i \rangle \langle y_j, \psi_k \rangle \langle \tilde{\psi}, \psi_i \rangle \langle \tilde{\psi}, \psi_k \rangle \right) \\ &= \sum_{i=1}^m \sum_{k=1}^m \left(\underbrace{\left\langle \sum_{j=1}^n \langle y_j, \psi_i \rangle y_j, \psi_k \right\rangle}_{=\lambda_i \psi_i} \langle \tilde{\psi}, \psi_i \rangle \langle \tilde{\psi}, \psi_k \rangle \right) \\ &= \sum_{i=1}^m \sum_{k=1}^m \left(\underbrace{\langle \lambda_i \psi_i, \psi_k \rangle}_{=\lambda_i \delta_{ik}} \langle \tilde{\psi}, \psi_i \rangle \langle \tilde{\psi}, \psi_k \rangle \right) \\ &= \sum_{i=1}^m \lambda_i |\langle \tilde{\psi}, \psi_i \rangle|^2 \leq \lambda_1 \sum_{i=1}^m |\langle \tilde{\psi}, \psi_i \rangle|^2 = \lambda_1 \|\tilde{\psi}\|^2 = \lambda_1 \\ &= \sum_{j=1}^n |\langle y_j, \psi_1 \rangle|^2. \end{aligned}$$

Luego, ψ_1 resuelve (\mathbf{P}^1) y $\arg \max(\mathbf{P}^1) = \sigma_1^2 = \lambda_1$.

Si ahora buscamos un segundo vector, ortogonal a ψ_1 que describa nuevamente el conjunto de datos $\{y_i\}_{i=1}^n$ con el menor error posible, entonces tenemos que resolver el siguiente problema:

$$(\mathbf{P}^2) \quad \max_{\tilde{\psi} \in \mathbb{R}^m} \sum_{j=1}^n |\langle y_j, \tilde{\psi} \rangle|^2 \quad \text{sueto a} \quad \|\tilde{\psi}\|^2 = 1 \quad \text{y} \quad \langle \tilde{\psi}, \psi_1 \rangle = 0.$$

La SVD implica que ψ_2 es una solución de (\mathbf{P}^2) y $\arg \max(\mathbf{P}^2) = \sigma_2^2 = \lambda_2$. De hecho, ψ_2 resuelve las condiciones necesarias de optimalidad de primer orden y para

$$\tilde{\psi} = \sum_{i=2}^m \langle \tilde{\psi}, \psi_i \rangle \psi_i \in \langle \psi_1 \rangle^\perp$$

tenemos

$$\sum_{j=1}^n |\langle y_j, \tilde{\psi} \rangle|^2 \leq \lambda_2 = \sum_{j=1}^n |\langle y_j, \psi_2 \rangle|^2.$$

Claramente, este procedimiento se puede continuar por inducción finita. Vamos a resumir los resultados obtenidos con el siguiente teorema.

Teorema 2.3.1. Sea $Y = (y_1 | \cdots | y_n) \in \mathbb{R}^{m \times n}$ una matriz de rango $d \leq \min\{m, n\}$ y denotemos por $Y = \Psi \Sigma \Phi^\top$ la descomposición en valores singulares de Y , donde $\Psi = (\psi_1 | \cdots | \psi_m) \in \mathbb{R}^{m \times m}$, $\Phi = (\phi_1 | \cdots | \phi_n) \in \mathbb{R}^{n \times n}$ son matrices ortogonales y la matriz $\Sigma \in \mathbb{R}^{m \times n}$ es:

$$\Sigma = \begin{pmatrix} D & \Theta \\ \Theta & \Theta \end{pmatrix},$$

donde $D = \text{diag}(\sigma_1, \dots, \sigma_d) \in \mathbb{R}^{d \times d}$. Entonces, para cualquier $\ell \in \{1, \dots, d\}$, la solución de

$$(\mathbf{P}^\ell) \quad \max_{\tilde{\psi}_1, \dots, \tilde{\psi}_\ell \in \mathbb{R}^m} \sum_{i=1}^{\ell} \sum_{j=1}^n |\langle y_j, \tilde{\psi}_i \rangle|^2 \quad \text{sujeto a} \quad \langle \tilde{\psi}_i, \tilde{\psi}_j \rangle = \delta_{ij} \text{ para } 1 \leq i, j \leq \ell$$

viene dada por los vectores $\{\psi_i\}_{i=1}^{\ell}$, es decir, por las primeras ℓ columnas independientes de Ψ , donde δ_{ij} es la delta de Kronecker. Además,

$$\arg \max (\mathbf{P}^\ell) = \sum_{i=1}^{\ell} \sigma_i^2 = \sum_{i=1}^{\ell} \lambda_i.$$

Antes de demostrar el anterior resultado, vamos a dar la siguiente definición:

Definición 2.3.2. Para $\ell \in \{1, \dots, d\}$, los vectores $\{\psi_i\}_{i=1}^{\ell}$ se llaman *base POD de rango ℓ* .

Demostración del Teorema. Como (\mathbf{P}^ℓ) es un problema de optimización con restricción de igualdad vamos a definir la función Lagrangiana

$$\mathcal{L} : \underbrace{\mathbb{R}^m \times \dots \times \mathbb{R}^m}_{\ell\text{-veces}} \times \mathbb{R}^{\ell \times \ell}$$

por

$$\mathcal{L}(\psi_1, \dots, \psi_\ell, \Lambda) = \sum_{i=1}^{\ell} \sum_{j=1}^n |\langle y_j, \psi_i \rangle|^2 + \sum_{i,j=1}^{\ell} \lambda_{ij} (\delta_{ij} - \langle \psi_i, \psi_j \rangle)$$

para $\psi_1, \dots, \psi_\ell \in \mathbb{R}^m$ y $\Lambda = ((\lambda_{ij})) \in \mathbb{R}^{\ell \times \ell}$. Las condiciones necesarias de optimalidad de primer orden para (\mathbf{P}^ℓ) vienen dadas por

$$\frac{\partial \mathcal{L}}{\partial \psi_k}(\psi_1, \dots, \psi_\ell, \Lambda) z = 0 \quad \text{para todo } z \in \mathbb{R}^m \text{ y } k \in \{1, \dots, \ell\}. \quad (2.5)$$

Ahora bien,

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \psi_k}(\psi_1, \dots, \psi_\ell, \Lambda) z &= 2 \sum_{i=1}^{\ell} \sum_{j=1}^n \langle y_j, \psi_i \rangle \langle y_j, z \rangle \delta_{ik} \\ &\quad - \sum_{i,j=1}^{\ell} \lambda_{ij} \langle \psi_i, z \rangle \delta_{jk} - \sum_{i,j=1}^{\ell} \lambda_{ij} \langle z, \psi_j \rangle \delta_{ki} \\ &= 2 \sum_{j=1}^n \langle y_j, \psi_k \rangle \langle y_j, z \rangle - \sum_{i=1}^{\ell} (\lambda_{ik} + \lambda_{ki}) \langle \psi_i, z \rangle \\ &= \left\langle 2 \sum_{j=1}^n \langle y_j, \psi_k \rangle y_j - \sum_{i=1}^{\ell} (\lambda_{ik} + \lambda_{ki}) \psi_i, z \right\rangle \end{aligned}$$

Luego (2.5), es equivalente a:

$$\sum_{j=1}^n \langle y_j, \psi_k \rangle y_j = \frac{1}{2} \sum_{i=1}^{\ell} (\lambda_{ik} + \lambda_{ki}) \psi_i \quad \text{en } \mathbb{R}^m \text{ para todo } k \in \{1, \dots, \ell\}. \quad (2.6)$$

Ahora bien, como

$$YY^\top \psi = \sum_{j=1}^n \langle y_j, \psi \rangle y_j \quad \text{para } \psi \in \mathbb{R}^m,$$

la condición (2.6) puede ser escrita de la forma:

$$YY^\top \psi_k = \frac{1}{2} \sum_{i=1}^{\ell} (\lambda_{ik} + \lambda_{ki}) \psi_i \quad \text{en } \mathbb{R}^m \text{ para todo } k \in \{1, \dots, \ell\}. \quad (2.7)$$

Ahora procedemos por inducción. Para $\ell = 1$ se escribe como

$$YY^\top \psi_1 = \lambda_1 \psi_1 \quad \text{en } \mathbb{R}^m$$

con $\lambda_1 = \lambda_{11}$. Ahora supongamos que para $\ell \geq 1$ las condiciones de optimalidad de primer orden vienen dadas por

$$YY^\top \psi_k = \lambda_k \psi_k \quad \text{en } \mathbb{R}^m \text{ para todo } k \in \{1, \dots, \ell\}. \quad (2.8)$$

Probemos para $\ell + 1$:

$$YY^\top \psi_{\ell+1} = \lambda_{\ell+1} \psi_{\ell+1} \quad \text{en } \mathbb{R}^m. \quad (2.9)$$

Como consecuencia de (2.7), se tiene que

$$YY^\top \psi_{\ell+1} = \frac{1}{2} \sum_{i=1}^{\ell+1} (\lambda_{i,\ell+1} + \lambda_{\ell+1,i}) \psi_i \quad \text{en } \mathbb{R}^m. \quad (2.10)$$

Como $\{\psi_i\}_{i=1}^{\ell+1}$ es una base POD, se cumple que $\langle \psi_{\ell+1}, \psi_j \rangle = 0$ para $1 \leq j \leq \ell$. Usando (2.8) y la simetría de la matriz YY^\top tenemos que para cualquier $j \in \{1, \dots, \ell\}$

$$\begin{aligned} 0 &= \lambda_j \langle \psi_{\ell+1}, \psi_j \rangle = \langle \psi_{\ell+1}, YY^\top \psi_j \rangle = \langle YY^\top \psi_{\ell+1}, \psi_j \rangle \\ &= \frac{1}{2} \sum_{i=1}^{\ell+1} (\lambda_{i,\ell+1} + \lambda_{\ell+1,i}) \langle \psi_i, \psi_j \rangle = (\lambda_{j,\ell+1} + \lambda_{\ell+1,j}). \end{aligned}$$

Luego,

$$\lambda_{\ell+1,i} = -\lambda_{i,\ell+1} \quad \text{para cualquier } i \in \{1, \dots, \ell\}. \quad (2.11)$$

Volviendo a (2.10) y utilizando (2.11), obtenemos:

$$\begin{aligned} YY^\top \psi_{\ell+1} &= \frac{1}{2} \sum_{i=1}^{\ell} (\lambda_{i,\ell+1} + \lambda_{\ell+1,i}) \psi_i + \lambda_{\ell+1,\ell+1} \psi_{\ell+1} \\ &= \frac{1}{2} \sum_{i=1}^{\ell} (\lambda_{i,\ell+1} - \lambda_{i,\ell+1}) \psi_i + \lambda_{\ell+1,\ell+1} \psi_{\ell+1} = \lambda_{\ell+1,\ell+1} \psi_{\ell+1}. \end{aligned}$$

Poniendo $\lambda_{\ell+1} = \lambda_{\ell+1,\ell+1}$ obtenemos (2.9).

Resumiendo, las condiciones necesarias de optimalidad para (\mathbf{P}^ℓ) vienen dadas por el problema de autovalores siguiente:

$$YY^\top \psi_i = \lambda_i \psi_i \quad \text{para } i = 1, \dots, \ell.$$

Se sigue de la SVD que $\{\psi_i\}_{i=1}^{\ell}$ resuelve el anterior problema de autovalores. La prueba de que $\{\psi_i\}_{i=1}^{\ell}$ es una solución de (\mathbf{P}^ℓ) y que $\arg\max(\mathbf{P}^\ell) = \sum_{i=1}^{\ell} \sigma_i^2$ se hace de manera análoga a la prueba para (\mathbf{P}^1) . \square

2.4. Propiedades de la base POD

El siguiente resultado establece que para cada $\ell \leq d$, la aproximación de las columnas de Y por los primeros vectores singulares $\{\psi_i\}_{i=1}^\ell$ es la óptima en cierto sentido que veremos a continuación.

Teorema 2.4.1 (Optimalidad de la base POD). Supongamos que estamos bajo las hipótesis del Teorema 2.3.1. Supongamos además que $\hat{\Psi}^d \in \mathbb{R}^{m \times d}$ es una matriz con vectores $\hat{\psi}_i$ dos a dos ortonormales y que la expansión de las columnas de Y en la base $\{\hat{\psi}_i\}_{i=1}^d$ viene dada por

$$Y = \hat{\Psi}^d C^d, \quad \text{donde } C_{ij}^d = \langle \hat{\psi}_i, y_j \rangle \text{ para } 1 \leq i \leq d; 1 \leq j \leq n.$$

Entonces, para cada $l \in \{1, \dots, d\}$, tenemos

$$\|Y - \Psi^\ell B^\ell\|_F \leq \|Y - \hat{\Psi}^\ell C^\ell\|_F. \quad (2.12)$$

En la expresión anterior, $\|\cdot\|_F$ denota la norma de Frobenius dada por:

$$\|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |A_{ij}|^2} = \sqrt{\text{Traza}(A^\top A)} \quad \text{para } A \in \mathbb{R}^{m \times n},$$

la matriz Ψ^ℓ denota las primeras $\ell \leq d$ columnas de Ψ , B^ℓ las primeras ℓ filas de B y análogamente se definen $\hat{\Psi}^\ell$ y C^ℓ .

Antes de dar la prueba de este resultado, vamos a dar un lema que nos establece que la norma de Frobenius es invariante bajo la multiplicación por matrices ortogonales.

Lema 2.4.2. Sea $A \in \mathbb{R}^{m \times n}$ y Q una matriz ortogonal de orden m . Entonces:

$$\|QA\|_F = \|A\|_F.$$

Demostración.

$$\|QA\|_F^2 = \text{Traza}((QA)^\top(QA)) = \text{Traza}(A^\top Q^\top QA) = \text{Traza}(A^\top A) = \|A\|_F^2.$$

□

Veamos ahora la prueba del Teorema 2.4.1

Demostración. Usando el Lema 2.4.2, es fácil ver que

$$\|Y - \hat{\Psi}^\ell C^\ell\|_F^2 = \|\hat{\Psi}^d(C^d - C_0^\ell)\|_F^2 = \|C^d - C_0^\ell\|_F^2 = \sum_{i=\ell+1}^d \sum_{j=1}^n |C_{ij}^d|^2,$$

donde $C_0^\ell \in \mathbb{R}^{d \times n}$ se obtiene a partir de $C \in \mathbb{R}^{d \times n}$ reemplazando las últimas $d - \ell$ filas por 0.

Análogamente,

$$\begin{aligned} \|Y - \Psi^\ell B^\ell\|_F^2 &= \|\Psi^k(B^d - B_0^\ell)\|_F^2 = \|B^d - B_0^\ell\|_F^2 = \sum_{i=\ell+1}^d \sum_{j=1}^n |B_{ij}^d|^2 \\ &= \sum_{i=\ell+1}^d \sum_{j=1}^n |\langle y_j, \psi_i \rangle|^2 = \sum_{i=\ell+1}^d \sum_{j=1}^n \langle \langle y_j, \psi_i \rangle y_j, \psi_i \rangle \\ &= \sum_{i=\ell+1}^d \langle Y Y^\top \psi_i, \psi_i \rangle = \sum_{i=\ell+1}^d \sigma_i^2 \end{aligned} \quad (2.13)$$

Por el Teorema 2.3.1, los vectores ψ_1, \dots, ψ_ℓ resuelven (\mathbf{P}^ℓ) . Además, usando (2.13),

$$\|Y\|_F^2 = \|\hat{\Psi}^d C^d\|_F^2 = \|C^d\|_F^2 = \sum_{i=1}^d \sum_{j=1}^n |C_{ij}^d|^2$$

y

$$\|Y\|_F^2 = \|\Psi^d B^d\|_F^2 = \|B^d\|_F^2 = \sum_{i=1}^d \sum_{j=1}^n |B_{ij}^d|^2 = \sum_{i=1}^d \sigma_i^2$$

se concluye que

$$\begin{aligned} \|Y - \Psi^\ell B^\ell\|_F^2 &= \sum_{i=\ell+1}^d \sigma_i^2 = \sum_{i=1}^d \sigma_i^2 - \sum_{i=1}^{\ell} \sigma_i^2 = \|Y\|_F^2 - \sum_{i=1}^{\ell} \sum_{j=1}^n |\langle y_j, \psi_i \rangle|^2 \\ &\leq \|Y\|_F^2 - \sum_{i=1}^{\ell} \sum_{j=1}^n |\langle y_j, \hat{\psi}_i \rangle|^2 = \sum_{i=1}^d \sum_{j=1}^n |C_{ij}^d|^2 - \sum_{i=1}^{\ell} \sum_{j=1}^n |C_{ij}^d|^2 \\ &= \sum_{i=\ell+1}^d \sum_{j=1}^n |C_{ij}^d|^2 = \|Y - \hat{\Psi}^\ell C^\ell\|_F^2 \end{aligned}$$

lo cual implica (2.12), como queríamos demostrar. \square

Observemos que

$$\begin{aligned} \|Y - \hat{\Psi}^\ell C^\ell\|_F^2 &= \sum_{i=1}^m \sum_{j=1}^n \left| Y_{ij} - \sum_{k=1}^{\ell} \hat{\Psi}_{ik}^\ell C_{kj} \right|^2 = \sum_{j=1}^n \sum_{i=1}^m \left| Y_{ij} - \sum_{k=1}^{\ell} \langle \hat{\psi}_k, y_j \rangle \hat{\Psi}_{ik}^\ell \right|^2 \\ &= \sum_{j=1}^n \left\| y_j - \sum_{k=1}^{\ell} \langle y_j, \hat{\psi}_k \rangle \hat{\psi}_k \right\|^2. \end{aligned}$$

Análogamente,

$$\|Y - \Psi^\ell B^\ell\|_F^2 = \sum_{j=1}^n \left\| y_j - \sum_{k=1}^{\ell} \langle y_j, \psi_k \rangle \psi_k \right\|^2.$$

Por tanto, (2.12) implica que

$$\sum_{j=1}^n \left\| y_j - \sum_{k=1}^{\ell} \langle y_j, \psi_k \rangle \psi_k \right\|^2 \leq \sum_{j=1}^n \left\| y_j - \sum_{k=1}^{\ell} \langle y_j, \hat{\psi}_k \rangle \hat{\psi}_k \right\|^2$$

para cualquier otro conjunto $\{\hat{\psi}_i\}_{i=1}^{\ell}$ de ℓ vectores dos a dos ortonormales. Luego, por el Teorema 2.4.1 la base POD de rango ℓ puede también ser determinada resolviendo el siguiente problema de optimización:

$$(\tilde{\mathbf{P}}^\ell) \quad \min_{\tilde{\psi}_1, \dots, \tilde{\psi}_\ell \in \mathbb{R}^m} \sum_{j=1}^n \left\| y_j - \sum_{i=1}^{\ell} \langle y_j, \tilde{\psi}_i \rangle \tilde{\psi}_i \right\|^2 \quad \text{sujeto a} \quad \langle \tilde{\psi}_i, \tilde{\psi}_j \rangle = \delta_{ij} \text{ para } 1 \leq i, j \leq \ell.$$

Nota 2.4.3 (Comparación de las condiciones de optimalidad entre (\mathbf{P}^ℓ) y $(\tilde{\mathbf{P}}^\ell)$).

Sea $\{\psi_i\}_{i=1}^{\ell}$ un conjunto de vectores ortonormales en \mathbb{R}^m , es decir,

$$\langle \psi_i, \psi_k \rangle = \delta_{ik} \quad \text{para } i, k \in \{1, \dots, m\}. \quad (2.14)$$

Para cualquier índice $k \in \{1, \dots, \ell\}$ y dirección $z \in \mathbb{R}^m$ tenemos:

$$\begin{aligned} 0 &= \frac{\partial}{\partial \psi_k} (\delta_{ik}) z = \frac{\partial}{\partial \psi_k} (\langle \psi_i, \psi_k \rangle) z \\ &= \begin{cases} \langle \psi_i, z \rangle & \text{para } i \in \{1, \dots, \ell\} \setminus \{k\} \\ 2 \langle \psi_i, z \rangle & \text{para } i = k \end{cases} \end{aligned}$$

Por tanto,

$$\langle \psi_i, z \rangle = 0 \quad \text{para } i \in \{1, \dots, \ell\} \text{ y } z \in \mathbb{R}^m. \quad (2.15)$$

Para $\ell \in \{1, \dots, m\}$, tomemos

$$z_j = z_j(\psi_1, \dots, \psi_\ell) = y_j - \sum_{i=1}^{\ell} \langle y_j, \psi_i \rangle \psi_i \in \mathbb{R}^m \quad \text{para } j = 1, \dots, n.$$

Sea

$$J(\psi_1, \dots, \psi_\ell) = \sum_{j=1}^n \|z_j\|^2.$$

Usando (2.14) obtenemos:

$$\begin{aligned} \|z_j\|_2^2 &= \langle z_j, z_j \rangle = \left\langle y_j - \sum_{k=1}^{\ell} \langle y_j, \psi_k \rangle \psi_k, y_j - \sum_{i=1}^{\ell} \langle y_j, \psi_i \rangle \psi_i \right\rangle \\ &= \langle y_j, y_j \rangle - 2 \sum_{i=1}^{\ell} \langle y_j, \psi_i \rangle \langle y_j, \psi_i \rangle + \sum_{i=1}^{\ell} \sum_{k=1}^{\ell} \langle y_j, \psi_i \rangle \langle y_j, \psi_k \rangle \langle \psi_i, \psi_k \rangle \\ &= \|y_j\|^2 - 2 \sum_{i=1}^{\ell} |\langle y_j, \psi_i \rangle|^2 + \sum_{i=1}^{\ell} |\langle y_j, \psi_i \rangle|^2 \\ &= \|y_j\|^2 - \sum_{i=1}^{\ell} |\langle y_j, \psi_i \rangle|^2. \end{aligned}$$

Luego:

$$J(\psi_1, \dots, \psi_\ell) = \sum_{j=1}^n \|z_j\|^2 = \sum_{j=1}^n \left(\|y_j\|^2 - \sum_{i=1}^{\ell} |\langle y_j, \psi_i \rangle|^2 \right) \quad (2.16)$$

Para cualquier $k \in \{1, \dots, \ell\}$ vamos a considerar las siguientes derivadas:

$$\frac{\partial}{\partial \psi_k} \left(\sum_{j=1}^n \left(\|y_j\|^2 - \sum_{i=1}^{\ell} |\langle y_j, \psi_i \rangle|^2 \right) \right) \quad \text{y} \quad \frac{\partial}{\partial \psi_k} \left(\sum_{j=1}^n \|z_j(\psi_1, \dots, \psi_\ell)\|^2 \right)$$

Debido a (2.16), ambas derivadas deben ser la misma. Observemos que

$$\begin{aligned} \frac{\partial J}{\partial \psi_k}(\psi_1, \dots, \psi_\ell) z &= \frac{\partial}{\partial \psi_k} \left(\sum_{j=1}^n \left(\|y_j\|^2 - \sum_{i=1}^{\ell} |\langle y_j, \psi_i \rangle|^2 \right) \right) z \\ &= \sum_{j=1}^n \frac{\partial}{\partial \psi_k} \left(\|y_j\|^2 - \sum_{i=1}^{\ell} |\langle y_j, \psi_i \rangle|^2 \right) z \\ &= - \sum_{j=1}^n 2 \langle y_j, \psi_k \rangle \langle y_j, z \rangle \\ &= \left\langle -2 \sum_{j=1}^n \langle y_j, \psi_k \rangle y_j, z \right\rangle \end{aligned}$$

para cualquier dirección $z \in \mathbb{R}^m$ y para $1 \leq k \leq \ell$. Notemos que

$$\sum_{j=1}^n \langle y_j, \psi \rangle y_j = YY^\top \psi \quad \text{para } \psi \in \mathbb{R}^m.$$

Por tanto, se tiene

$$\frac{\partial J}{\partial \psi_k}(\psi_1, \dots, \psi_\ell) = -2YY^\top \psi_k \quad \text{para } 1 \leq k \leq \ell. \quad (2.17)$$

Por otro lado tenemos

$$\frac{\partial z_j}{\partial \psi_k} z = -\langle y_j, \psi_k \rangle z - \langle y_j, z \rangle \psi_k = -\langle y_j, \psi_k \rangle z - \langle y_j, z \rangle \psi_k$$

para $1 \leq k \leq \ell$ y $z \in \mathbb{R}^m$. Ahora, usando (2.14) y (2.15) obtenemos:

$$\begin{aligned} \frac{\partial}{\partial \psi_k} (\|z_j\|^2) z &= \frac{\partial}{\partial \psi_k} (\langle z_j, z_j \rangle) z = 2 \left\langle z_j, \frac{\partial z_j}{\partial \psi_k} z \right\rangle \\ &= 2 \left\langle y_j - \sum_{i=1}^{\ell} \langle y_j, \psi_i \rangle \psi_i, -\langle y_j, z \rangle \psi_k - \langle y_j, \psi_k \rangle z \right\rangle \\ &= -2 \langle y_j, z \rangle \langle y_j, \psi_k \rangle_{\mathbb{R}^m} + 2 \sum_{i=1}^{\ell} \langle y_j, \psi_i \rangle \langle y_j, z \rangle \langle y_j, \psi_k \rangle \\ &\quad - 2 \langle y_j, \psi_k \rangle \langle y_j, z \rangle_{\mathbb{R}^m} + 2 \sum_{i=1}^{\ell} \langle y_j, \psi_i \rangle \langle y_j, \psi_k \rangle \langle y_j, z \rangle \\ &= -2 \langle y_j, \psi_k \rangle \langle y_j, z \rangle = \langle -2 \langle y_j, \psi_k \rangle y_j, z \rangle \end{aligned}$$

para cualquier dirección $z \in \mathbb{R}^m$, $j = 1, \dots, n$ y $1 \leq k \leq \ell$.

Resumiendo, tenemos

$$\frac{\partial J}{\partial \psi_k}(\psi_1, \dots, \psi_\ell) = -2 \sum_{j=1}^n \langle y_j, \psi_k \rangle y_j = -2YY^\top \psi_k$$

el cual coincide con (2.17).

La siguiente consecuencia del Teorema 2.3.1 nos establece que los coeficientes POD están incorrelados.

Corolario 2.4.4. Supongamos que estamos bajo las hipótesis del Teorema 2.3.1.

Entonces:

$$\sum_{j=1}^n \langle y_j, \psi_i \rangle \langle y_j, \psi_k \rangle = \sum_{j=1}^n B_{ij}^\ell B_{kj}^\ell = \sigma_i^2 \delta_{ik} \quad \text{para } i, k \in \{1, \dots, \ell\}.$$

Demostración. La prueba es una consecuencia inmediata de (2.14):

$$\sum_{j=1}^n \langle y_j, \psi_i \rangle \langle y_j, \psi_k \rangle = \left\langle \underbrace{\sum_{j=1}^n \langle y_j, \psi_i \rangle y_j}_{=YY^\top \psi_i}, \psi_k \right\rangle = \langle \sigma_i^2 \psi_i, \psi_k \rangle = \sigma_i^2 \delta_{ik}.$$

□

Cálculo efectivo de la base POD

Vamos a terminar con una observación sobre la computación práctica de una base POD de rango ℓ .

- Si $n < m$, entonces se puede determinar la base POD de rango ℓ de la siguiente forma: Se calculan los autovectores $\phi_1, \dots, \phi_\ell \in \mathbb{R}^n$ resolviendo el problema de autovalores $n \times n$:

$$Y^\top Y \phi_i = \lambda_i \phi_i \quad \text{para } i = 1, \dots, \ell \quad (2.18)$$

y se toma por (2.2),

$$\psi_i = \frac{1}{\sqrt{\lambda_i}} Y \phi_i \quad \text{para } i = 1, \dots, \ell. \quad (2.19)$$

- Si $n > m$, entonces se puede calcular la base POD resolviendo el problema de autovalores $m \times m$:

$$YY^\top \psi_i = \lambda_i \psi_i \quad \text{para } i = 1, \dots, \ell. \quad (2.20)$$

Por razones históricas [11] este método de determinar la base POD es a menudo llamado *método de los snapshots*.

Para la aplicación de este método POD a problemas concretos, la elección de ℓ es de vital importancia. En general, no hay reglas definidas para ello. Más bien, su elección se basa en consideraciones heurísticas combinadas con la observación de la relación entre energía modelada y energía total contenida en el sistema Y , lo cual viene expresado como

$$\mathcal{E}(\ell) = \frac{\sum_{i=1}^{\ell} \lambda_i}{\sum_{i=1}^d \lambda_i}. \quad (2.21)$$

Notemos que $\sum_{i=1}^{\ell} \lambda_i = \text{Traza}(YY^\top) = \text{Traza}(Y^\top Y)$. Mencionemos además que el método POD se llama también *Análisis de Componentes Principales* o *Descomposición de Karhunen-Loève*.

2.5. El método POD para productos escalares con peso

Dotemos a \mathbb{R}^m del siguiente producto escalar con peso:

$$\langle \psi, \tilde{\psi} \rangle_W = \psi^\top W \tilde{\psi} = \langle \psi, W \tilde{\psi} \rangle_{\mathbb{R}^m} = \langle W \psi, \tilde{\psi} \rangle_{\mathbb{R}^m} \quad \text{para } \psi, \tilde{\psi} \in \mathbb{R}^m, \quad (2.22)$$

donde $W \in \mathbb{R}^{m \times m}$ es una matriz simétrica y definida positiva. Además, sea $\|\psi\|_W = \sqrt{\langle \psi, \psi \rangle_W}$ para $\psi \in \mathbb{R}^m$, la norma inducida asociada. Observemos que para la elección de $W = I_m$, el producto escalar anteriormente definido coincide con el producto escalar euclídeo.

Motivemos esta definición con el siguiente ejemplo.

Ejemplo 2.5.1. Sea $L^2(0, 1)$ el espacio de las funciones medibles de cuadrado integrable en $(0, 1)$, es decir,

$$L^2(0, 1) = \left\{ \varphi : (0, 1) \rightarrow \mathbb{R} \text{ medibles} : \int_0^1 |\varphi|^2 dx < +\infty \right\}.$$

Es bien conocido que $L^2(0, 1)$ es un espacio de Hilbert con el producto escalar

$$\langle \varphi, \tilde{\varphi} \rangle_{L^2(0,1)} = \int_0^1 \varphi \tilde{\varphi} dx \quad \text{para } \varphi, \tilde{\varphi} \in L^2(0, 1)$$

cuya norma inducida es $\|\varphi\|_{L^2(0,1)} = \sqrt{\langle \varphi, \varphi \rangle_{L^2(0,1)}}$ para $\varphi \in L^2(0, 1)$.

Nuestro objetivo es mostrar que el producto escalar en $L^2(0, 1)$, espacio de dimensión infinita, se puede aproximar por un cierto producto escalar discreto en un espacio de dimensión finita. Para ello, consideramos una partición uniforme del intervalo $(0, 1)$:

$$\{0 = x_1 < x_2 < \cdots < x_{m-1} < x_m = 1\},$$

donde

$$x_j = (j - 1)h \quad \text{para } j = 1, \dots, m,$$

$$h = 1/(m - 1).$$

Vamos a aproximar la integral $\int_0^1 \varphi(x)\tilde{\varphi}(x) dx$ usando la fórmula de cuadratura del trapecio:

$$\begin{aligned} \langle \varphi, \tilde{\varphi} \rangle_{L^2(0,1)} &= \int_0^1 \varphi(x)\tilde{\varphi}(x) dx = \sum_{j=1}^{m-1} \int_{x_j}^{x_{j+1}} \varphi(x)\tilde{\varphi}(x) dx \\ &\approx \frac{h}{2}\varphi(x_1)\tilde{\varphi}(x_1) + h \sum_{j=2}^{m-1} \varphi(x_j)\tilde{\varphi}(x_j) + \frac{h}{2}\varphi(x_m)\tilde{\varphi}(x_m). \end{aligned}$$

Si tomamos $\varphi^h = (\varphi_1^h, \dots, \varphi_m^h)^\top \in \mathbb{R}^m$ y $\tilde{\varphi}^h = (\tilde{\varphi}_1^h, \dots, \tilde{\varphi}_m^h)^\top \in \mathbb{R}^m$, donde $\varphi_i^h = \varphi(x_i)$ y $\tilde{\varphi}_i^h = \tilde{\varphi}(x_i)$ para todo $j = 1, \dots, m$, definimos el siguiente producto escalar discreto en $L^2(0, 1)$:

$$\langle \varphi, \tilde{\varphi} \rangle_{L_h^2(0,1)} := h \left(\frac{\varphi_1^h \tilde{\varphi}_1^h}{2} + \sum_{j=2}^{m-1} (\varphi_j^h \tilde{\varphi}_j^h) + \frac{\varphi_m^h \tilde{\varphi}_m^h}{2} \right). \quad (2.23)$$

Si escribimos (2.23) en forma matricial, tenemos:

$$\langle \varphi, \tilde{\varphi} \rangle_{L_h^2(0,1)} = \begin{pmatrix} \varphi_1^h & \dots & \varphi_m^h \end{pmatrix} W \begin{pmatrix} \tilde{\varphi}_1^h \\ \vdots \\ \tilde{\varphi}_m^h \end{pmatrix} = \varphi^h W \tilde{\varphi}^h,$$

donde $W := \text{diag}(h/2, h, \dots, h, h/2) \in \mathbb{R}^{m \times m}$. Es decir, hemos escrito el producto escalar de la misma forma que aparecía en (2.22):

$$\langle \varphi, \tilde{\varphi} \rangle_{L_h^2(0,1)} = (\varphi^h)^\top W \tilde{\varphi}^h = \langle \varphi^h, \tilde{\varphi}^h \rangle_W.$$

Ahora vamos a reemplazar (\mathbf{P}^1) por

$$(\mathbf{P}_W^1) \quad \max_{\psi \in \mathbb{R}^m} \sum_{j=1}^n |\langle y_j, \tilde{\psi} \rangle_W|^2 \quad \text{sujeto a} \quad \|\tilde{\psi}\|_W^2 = 1.$$

Análogamente a lo que hicimos en la Sección 2.3, vamos a tratar (\mathbf{P}_W^1) como un problema de optimización con restricción de igualdad. El Lagrangiano $\mathcal{L} : \mathbb{R}^m \times \mathbb{R} \rightarrow \mathbb{R}$ asociado a (\mathbf{P}_W^1) viene dado por

$$\mathcal{L}(\psi, \lambda) = \sum_{j=1}^n |\langle y_j, \psi \rangle_W|^2 + \lambda(1 - \|\psi\|_W^2) \quad \text{para} \quad (\psi, \lambda) \in \mathbb{R}^m \times \mathbb{R}.$$

Introducimos la función $e : \mathbb{R}^m \rightarrow \mathbb{R}$ definida por $e(\psi) = 1 - \|\psi\|_W^2 = 1 - \psi^\top W \psi$ para $\psi \in \mathbb{R}^m$. Entonces, la restricción del problema puede reescribirse como $e(\psi) = 0$. Observemos que cualquier solución ψ de (\mathbf{P}_W^1) es no nula, ya que tiene que verificar $\|\psi\|_W = 1$. Por tanto, para cualquier solución ψ del problema, el conjunto $\{\nabla e(\psi)\} = \{-2\psi^\top W\}$ es linealmente independiente y por ende cualquier solución de (\mathbf{P}_W^1) es un punto regular. En consecuencia, existe un único multiplicador de Lagrange $\lambda \in \mathbb{R}$ asociado con la solución óptima ψ verificando la condición necesaria de optimalidad de primer orden

$$\nabla \mathcal{L}(\psi, \lambda) = 0 \quad \text{en} \quad \mathbb{R}^m \times \mathbb{R}.$$

Vamos a empezar calculando el gradiente de \mathcal{L} con respecto a ψ . Como W es una matriz simétrica, se tiene:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \psi_i}(\psi, \lambda) &= \frac{\partial}{\partial \psi_i} \left(\sum_{j=1}^n \left| \sum_{k=1}^m \sum_{\nu=1}^m Y_{j\nu}^\top W_{\nu k} \psi_k \right|^2 + \lambda \left(1 - \sum_{k=1}^m \sum_{\nu=1}^m \psi_\nu W_{\nu k} \psi_k \right) \right) \\ &= 2 \sum_{j=1}^n \left(\sum_{k=1}^m \sum_{\nu=1}^m Y_{j\nu}^\top W_{\nu k} \psi_k \right) \left(\sum_{\mu=1}^m Y_{j\mu}^\top W_{\mu i} \right) \\ &\quad - \lambda \left(\sum_{\nu=1}^m \psi_\nu W_{\nu i} + \sum_{k=1}^m W_{ik} \psi_k \right) \\ &= 2 \sum_{k=1}^m \sum_{\nu=1}^m \sum_{\mu=1}^m W_{i\mu} \sum_{j=1}^n Y_{\mu j} Y_{j\nu}^\top W_{\nu k} \psi_k - 2\lambda \left(\sum_{k=1}^m W_{ik} \psi_k \right) \\ &= 2 \left(W Y Y^\top W \psi - \lambda W \psi \right)_i. \end{aligned}$$

Por tanto:

$$\nabla_\psi \mathcal{L}(\psi, \lambda) = 2 \left(W Y Y^\top W \psi - \lambda W \psi \right).$$

Si imponemos que $\nabla_\psi \mathcal{L}(\psi, \lambda) = 0$, entonces tenemos que resolver el siguiente problema de autovalores generalizado

$$(WY)(WY)^\top \psi = \lambda W \psi. \quad (2.24)$$

Como W es simétrica y definida positiva, posee una descomposición de la forma $W = QDQ^\top$, donde $D = \text{diag}(\eta_1, \dots, \eta_m)$ que contiene los autovalores $\eta_1 \geq \dots \geq \eta_m > 0$ de W y $Q \in \mathbb{R}^{m \times m}$ es una matriz ortogonal. Definamos:

$$W^\alpha := Q \text{diag}(\eta_1^\alpha, \dots, \eta_m^\alpha) Q^\top \quad \text{para } \alpha \in \mathbb{R}.$$

Es fácil comprobar las siguientes propiedades de la matriz W^α .

- $(W^\alpha)^{-1} = W^{-\alpha}$ para cualquier $\alpha \in \mathbb{R}$.
- $W^{\alpha+\beta} = W^\alpha W^\beta$ para cualesquiera $\alpha, \beta \in \mathbb{R}$.

Además, tenemos

$$\langle \psi, \tilde{\psi} \rangle_W = \langle W^{1/2}\psi, W^{1/2}\tilde{\psi} \rangle_{\mathbb{R}^m} \text{ y } \|\psi\|_W = \left\| W^{1/2}\psi \right\|_{\mathbb{R}^m} \text{ para } \psi, \tilde{\psi} \in \mathbb{R}^m.$$

Tomando $\bar{\psi} = W^{1/2}\psi \in \mathbb{R}^m$ e $\bar{Y} = W^{1/2}Y \in \mathbb{R}^{m \times n}$ y multiplicando (2.24) por $W^{-1/2}$ por la izquierda, llegamos al siguiente problema de autovalores

$$\bar{Y}\bar{Y}^\top \bar{\psi} = \lambda \bar{\psi} \text{ en } \mathbb{R}^m.$$

De $\frac{\partial \mathcal{L}}{\partial \lambda}(\psi, \lambda) = 0$ en \mathbb{R} obtenemos la restricción $\|\psi\|_W = 1$ que puede también ser expresada como

$$\|\bar{\psi}\|_{\mathbb{R}^m} = 1.$$

Por tanto, las condiciones de optimalidad de primer orden para (\mathbf{P}_W^1) son análogas a las de (\mathbf{P}^1) (2.4), pero en este caso, la matriz Y y el vector ψ tienen que estar pesados por la matriz $W^{1/2}$.

Notemos que $\nabla_{\psi\psi} \mathcal{L}(\psi, \lambda) = 2(WYY^\top W\psi - \lambda W\psi) \in \mathbb{R}^{m \times m}$. Sea $\psi \in \mathbb{R}^m$ un vector arbitrario. Como $\bar{Y}\bar{Y}^\top$ es simétrica, existen m autovectores ortonormales (con respecto al producto escalar euclídeo) $\bar{\psi}_1, \dots, \bar{\psi}_m \in \mathbb{R}^m$ de $\bar{Y}\bar{Y}^\top$ verificando $\bar{Y}\bar{Y}^\top \bar{\psi}_i = \lambda_i \bar{\psi}_i$ para $1 \leq i \leq m$. Tomemos $\psi_i = W^{-1/2}\bar{\psi}_i$, $1 \leq i \leq m$. Entonces, $\{\psi_i\}_{i=1}^m$ forman una base ortonormal (con respecto al producto escalar con peso) en \mathbb{R}^m y $WYY^\top W\psi_i = \lambda_i W\psi_i$. Podemos escribir ψ en la forma

$$\psi = \sum_{i=1}^m \langle \psi, \psi_i \rangle_{\mathbb{R}^m} \psi_i.$$

Veamos que $\nabla_{\psi\psi}\mathcal{L}(\psi_1, \lambda_1)$ es semidefinida negativa:

$$\begin{aligned}\psi^\top \nabla_{\psi\psi}\mathcal{L}(\psi_1, \lambda_1) \psi &= 2\psi^\top \left(WY Y^\top W - \lambda_1 W \right) \psi \\ &= 2 \sum_{i=1}^m \sum_{j=1}^m \langle \psi, \psi_i \rangle_{\mathbb{R}^m} \langle \psi, \psi_j \rangle_{\mathbb{R}^m} \psi_i^\top \left(WY Y^\top W - \lambda_1 W \right) \psi_j \\ &= 2 \sum_{i=1}^m \sum_{j=1}^m (\lambda_j - \lambda_1) \langle \psi, \psi_i \rangle_{\mathbb{R}^m} \langle \psi, \psi_j \rangle_{\mathbb{R}^m} \psi_i^\top W \psi_j \\ &= 2 \sum_{i=1}^m (\lambda_j - \lambda_1) |\langle \psi, \psi_i \rangle_{\mathbb{R}^m}|^2 \leq 0.\end{aligned}$$

Por tanto, la matriz $\nabla_{\psi\psi}\mathcal{L}(\psi_1, \lambda_1)$ es semidefinida negativa, lo cual implica que se verifica la condición necesaria de optimalidad de segundo orden. No obstante, haciendo un argumento análogo al de la Sección 2.3, se puede probar que

$$\psi_1 = W^{-1/2} \bar{\psi}_1$$

resuelve (\mathbf{P}_W^1) , donde $\bar{\psi}_1$ es un autovector de $\bar{Y}\bar{Y}^\top$ correspondiente al mayor autovalor λ_1 con $\|\bar{\psi}_1\|_{\mathbb{R}^m} = 1$. Debido a la descomposición en valores singulares (SVD), el vector ψ_1 puede también ser determinado resolviendo el problema de autovalores

$$\bar{Y}^\top \bar{Y} \bar{\phi}_1 = \lambda_1 \bar{\phi}_1$$

donde $\bar{Y}^\top \bar{Y} = Y^\top W Y$, y tomando

$$\psi_1 = W^{-1/2} \bar{\psi}_1 = \frac{1}{\sqrt{\lambda_1}} W^{-1/2} \bar{Y} \bar{\phi}_1 = \frac{1}{\sqrt{\lambda_1}} Y \bar{\phi}_1.$$

Como en la Sección 2.1, podemos continuar buscando un segundo vector $\psi \in \mathbb{R}^m$ con $\langle \psi, \psi_1 \rangle_W = 0$ que maximice $\sum_{j=1}^n |\langle y_j, \psi \rangle_W|^2$.

Generalicemos a continuación el Teorema 2.3.1:

Teorema 2.5.2. Sea $Y \in \mathbb{R}^{m \times n}$ una matriz con rango $d \leq \min\{m, n\}$, W una matriz simétrica y definida positiva, $\bar{Y} = W^{1/2} Y$ y $\ell \in \{1, \dots, d\}$. Además, sea $\bar{Y} = \bar{\Psi} \bar{\Sigma} \bar{\Phi}^\top$ la descomposición en valores singulares de \bar{Y} , donde $\bar{\Psi} = [\bar{\psi}_1, \dots, \bar{\psi}_m] \in \mathbb{R}^{m \times m}$, $\bar{\Phi} = [\bar{\phi}_1, \dots, \bar{\phi}_n] \in \mathbb{R}^{n \times n}$ son matrices ortogonales y $\bar{\Sigma}$ una matriz de la forma

$$\bar{\Psi}^\top \bar{Y} \bar{\Phi} = \begin{pmatrix} D & \Theta \\ \Theta & \Theta \end{pmatrix} = \Sigma \in \mathbb{R}^{m \times n}.$$

Entonces, la solución de

$$(\mathbf{P}_W^\ell) \quad \max_{\tilde{\psi}_1, \dots, \tilde{\psi}_\ell \in \mathbb{R}^m} \sum_{i=1}^{\ell} \sum_{j=1}^n |\langle y_j, \tilde{\psi}_i \rangle_W|^2 \quad \text{sujeto a} \quad \langle \tilde{\psi}_i, \tilde{\psi}_j \rangle_W = \delta_{ij} \text{ para } 1 \leq i, j \leq \ell$$

viene dada por los vectores $\psi_i = W^{-1/2} \bar{\psi}_i, i = 1, \dots, \ell$. Además,

$$\operatorname{argmax}(\mathbf{P}_W^\ell) = \sum_{i=1}^{\ell} \sigma_i^2 = \sum_{i=1}^{\ell} \lambda_i.$$

Demostración. Usando argumentos similares de la prueba del Teorema 2.3.1, se puede demostrar que $\{\psi_i\}_{i=1}^{\ell}$ resuelve (\mathbf{P}_W^ℓ) . \square

Cálculo efectivo de la base POD

Debido a la SVD y $\bar{Y}^\top \bar{Y} = Y^\top W Y$, la base POD $\{\psi_i\}_{i=1}^{\ell}$ de rango ℓ puede ser determinada por el método de los snapshots de la siguiente forma.

Resolver el problema de autovalores:

$$Y^\top W Y \bar{\phi}_i = \lambda_i \bar{\phi}_i \quad \text{para } i = 1, \dots, \ell$$

y tomar

$$\psi_i = W^{-1/2} \bar{\psi}_i = \frac{1}{\sqrt{\lambda_i}} W^{-1/2} (\bar{Y} \bar{\phi}_i) = \frac{1}{\sqrt{\lambda_i}} W^{-1/2} W^{1/2} Y \bar{\phi}_i = \frac{1}{\sqrt{\lambda_i}} Y \bar{\phi}_i$$

para $i = 1, \dots, \ell$. Notemos que

$$\langle \psi_i, \psi_j \rangle_W = \psi_i^\top W \psi_j = \frac{\delta_{ij} \lambda_j}{\sqrt{\lambda_i \lambda_j}} \quad \text{para } 1 \leq i, j \leq \ell.$$

Para $m \gg n$ el método de los snapshots resulta ser más rápido que calcular la base POD resolviendo el problema

$$\begin{cases} \bar{Y} \bar{Y}^\top \bar{\psi} = \lambda \bar{\psi} & \text{en } \mathbb{R}^m, \\ \|\bar{\psi}\|_{\mathbb{R}^m} = 1 & \text{en } \mathbb{R}. \end{cases}$$

Nota 2.5.3. Siguiendo las mismas ideas que las ya desarrolladas en la Sección 2.4 podemos dar una reformulación del problema (\mathbf{P}_W^ℓ) . Nos planteamos, entonces, resolver el siguiente problema de optimización:

$$(\hat{\mathbf{P}}_W^{n,\ell}) \quad \min_{\tilde{\psi}_1, \dots, \tilde{\psi}_\ell \in \mathbb{R}^m} \sum_{j=1}^n \alpha_j \left\| y_j - \sum_{i=1}^{\ell} \langle y_j, \tilde{\psi}_i \rangle_W \tilde{\psi}_i \right\|_W^2$$

s.t. $\langle \tilde{\psi}_i, \tilde{\psi}_j \rangle_W = \delta_{ij}$ para $1 \leq i, j \leq \ell$,

donde los α_j denotan pesos no negativos que especificaremos más adelante. Observe-
mos que si tomamos $\alpha_j = 1$ para todo $j = 1, \dots, n$ y $W = I_m$, el problema $(\hat{\mathbf{P}}_W^{n,\ell})$
coincide con $(\hat{\mathbf{P}}^\ell)$.

Para resolver $(\hat{\mathbf{P}}_W^{n,\ell})$ vamos a usar las técnicas que hemos visto con anterioridad.
Definimos la función Lagrangiana

$$\mathcal{L} : \underbrace{\mathbb{R}^m \times \dots \times \mathbb{R}^m}_{\ell \text{ veces}} \times \mathbb{R}^{\ell \times \ell} \rightarrow \mathbb{R}$$

por

$$\mathcal{L}(\psi_1, \dots, \psi_\ell, \Lambda) = \sum_{j=1}^n \alpha_j \left\| y_j - \sum_{i=1}^{\ell} \langle y_j, \psi_i \rangle_W \psi_i \right\|_W^2 + \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} \Lambda_{ij} (1 - \langle \psi_i, \psi_j \rangle_W)$$

para $\psi_1, \dots, \psi_\ell \in \mathbb{R}^m$ y $\Lambda \in \mathbb{R}^{\ell \times \ell}$ con elementos $\Lambda_{ij} = \delta_{ij}$, $1 \leq i, j \leq \ell$. Resulta que
la solución de $(\hat{\mathbf{P}}_W^{n,\ell})$ viene dada por las condiciones de optimalidad necesarias de
primer orden:

$$\nabla_{\psi_i} \mathcal{L}(\psi_1, \dots, \psi_\ell, \Lambda) = 0 \quad \text{en } \mathbb{R}^m, 1 \leq i \leq \ell \quad (2.25)$$

y

$$\langle \psi_i, \psi_j \rangle_W = \delta_{ij}, \quad 1 \leq i, j \leq \ell. \quad (2.26)$$

De (2.25), haciendo cálculos similares a los de la prueba del Teorema 2.3.1, se tiene:

$$YDY^\top W\psi_i = \lambda_i \psi_i \quad \text{para } i = 1, \dots, \ell \quad (2.27)$$

donde $D = \text{diag}(\alpha_1, \dots, \alpha_n) \in \mathbb{R}^{n \times n}$. Escribiendo $\psi_i = W^{-1/2} \bar{\psi}_i$ en (2.27) y multi-
plicando (2.27) por $W^{1/2}$ por la izquierda, llegamos a

$$W^{1/2} Y D Y^\top W^{1/2} \bar{\psi}_i = \lambda_i \bar{\psi}_i.$$

Por otro lado, de (2.26), obtenemos:

$$\langle \bar{\psi}_i, \bar{\psi}_j \rangle_{\mathbb{R}^m} = \bar{\psi}_i^\top \bar{\psi}_j = \psi_i^\top W \psi_j = \langle \psi_i, \psi_j \rangle_W = \delta_{ij}, \quad 1 \leq i, j \leq \ell.$$

Tomando $\bar{Y} = W^{1/2} Y D^{1/2} \in \mathbb{R}^{m \times n}$ y usando el hecho de que las matrices W y D
son simétricas, inferimos que la solución $\{\psi_i\}_{i=1}^{\ell}$ de $(\hat{\mathbf{P}}_W^{n,\ell})$ viene dada a través del
siguiente problema de autovalores simétrico $m \times m$:

$$\begin{cases} \bar{Y} \bar{Y}^\top \bar{\psi}_i = \lambda_i \bar{\psi}_i, & 1 \leq i \leq \ell, \\ \langle \bar{\psi}_i, \bar{\psi}_j \rangle_{\mathbb{R}^m} = \delta_{ij}, & 1 \leq i, j \leq \ell. \end{cases}$$

Notemos además que

$$\bar{Y}^\top \bar{Y} = D^{1/2} Y^\top W Y D^{1/2} \in \mathbb{R}^{n \times n}.$$

Por tanto, la base POD de rango ℓ puede también ser calculada por el método de los snapschots de la siguiente forma.

Cálculo efectivo de la base POD

Hay que resolver, en primer lugar, el problema de autovalores simétrico $n \times n$ dado por

$$\begin{cases} \bar{Y}^\top \bar{Y} \bar{\phi}_i = \lambda_i \bar{\phi}_i, & 1 \leq i \leq \ell \\ \langle \bar{\phi}_i, \bar{\phi}_j \rangle_{\mathbb{R}^n} = \delta_{ij}, & 1 \leq i, j \leq \ell \end{cases}$$

Después, por la SVD, tomamos

$$\psi_i = W^{-1/2} \bar{\psi}_i = \frac{1}{\sqrt{\lambda_i}} W^{-1/2} \bar{Y} \bar{\phi}_i = \frac{1}{\sqrt{\lambda_i}} Y D^{1/2} \bar{\phi}_i, \quad 1 \leq i \leq \ell.$$

Observemos que

$$\langle \psi_i, \psi_j \rangle_W = \psi_i^\top W \psi_j = \frac{1}{\sqrt{\lambda_i \lambda_j}} \bar{\psi}_i^\top \underbrace{D^{1/2} Y^\top W Y D^{1/2}}_{=\bar{Y}^\top \bar{Y}} \bar{\psi}_j = \frac{\lambda_i}{\sqrt{\lambda_i \lambda_j}} \bar{\phi}_i^\top \bar{\phi}_j = \frac{\lambda_i \delta_{ij}}{\sqrt{\lambda_i \lambda_j}}$$

para $1 \leq i, j \leq \ell$, es decir, los vectores de la base POD ψ_1, \dots, ψ_ℓ son ortonormales en \mathbb{R}^m con respecto al producto escalar $\langle \cdot, \cdot \rangle_W$.

2.6. Algunas aplicaciones del método POD

2.6.1. Ecuaciones diferenciales ordinarias

Para $T > 0$ vamos a considerar el siguiente problema de valor inicial semilineal:

$$(SL) \quad \begin{cases} \dot{y}(t) = Ay(t) + f(t, y(t)) & \text{para } t \in (0, T], \\ y(0) = y_0, \end{cases}$$

donde $y_0 \in \mathbb{R}^m$ es una condición inicial elegida, $A \in \mathbb{R}^{m \times m}$ es una matriz dada, $f : [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ es una función continua en ambos argumentos y Lipschitziana con respecto a la segunda variable.

Es bien conocido que el problema (SL) tiene una única solución clásica $y \in C^1(0, T; \mathbb{R}^m) \cap$

$\mathcal{C}([0, T]; \mathbb{R}^m)$. Entonces, la solución y de (SL) viene dada por la siguiente expresión integral en forma implícita:

$$y(t) = e^{tA}y_0 + \int_0^t e^{(t-s)A}f(s, y(s)) ds,$$

$$\text{con } e^{tA} = \sum_{n=0}^{\infty} \frac{t^n A^n}{n!}.$$

Vamos a considerar ahora una partición del intervalo $[0, T]$, es decir, $\{0 = t_1 < t_2 < \dots < t_n = T\}$. Por simplificación, vamos a escoger la partición de manera uniforme con paso de tiempo $\Delta t = T/(n-1)$, es decir, $t_j = (j-1)\Delta t$. Supongamos que conocemos la solución de (SL) en los instantes de tiempo t_j para $j = 1, \dots, n$. Nuestro objetivo es determinar una base POD de rango $\ell \leq \min\{m, n\}$ que describa el conjunto

$$y_j = y(t_j) = e^{t_j A}y_0 + \int_0^{t_j} e^{(t_j-s)A}f(s, y(s)) ds, \quad j = 1, \dots, n,$$

con el menor error posible. Para ello, nos planteamos el problema $(\hat{\mathbf{P}}_W^{n, \ell})$. Observemos que hemos pasado del caso continuo al discreto calculando $y_j = y(t_j)$ para cada $j = 1, \dots, n$.

El conjunto de snapshots $\{y_j\}_{j=1}^n$ para $(\hat{\mathbf{P}}_W^{n, \ell})$ y por tanto el conjunto generado por $\{y_1, \dots, y_n\}$ dependen de la elección de los instantes de tiempo $\{t_j\}_{j=1}^n$. Por consiguiente, los vectores de la base POD, $\{\psi_i\}_{i=1}^{\ell}$ y los correspondientes autovalores $\{\lambda_i\}_{i=1}^{\ell}$ dependen también de los instantes de tiempo, es decir,

$$\psi_i = \psi_i^n \quad \text{y} \quad \lambda_i = \lambda_i^n, \quad 1 \leq i \leq \ell.$$

Además, no hemos discutido hasta ahora cuál es la motivación para introducir los pesos no negativos $\{\alpha_j\}_{j=1}^n$ en $(\hat{\mathbf{P}}_W^{n, \ell})$. Por esta razón nos planteamos las siguientes cuestiones:

- ¿Cómo elegir unos instantes de tiempo que te permitan obtener una buena aproximación en otros tiempos?
- ¿Cuáles son los apropiados pesos no negativos $\{\alpha_j\}_{j=1}^n$ que debemos elegir para que el producto escalar con peso que genera estos valores $\{\alpha_j\}_{j=1}^n$ aproximen al producto escalar continuo?

Para responder a estas dos preguntas vamos a introducir una versión continua del método POD y supondremos que el problema (SL) tiene una única solución $y : [0, T] \rightarrow \mathbb{R}^m$. Si estamos interesados en encontrar una base POD de rango ℓ que describa toda la trayectoria $\{y(t) : t \in [0, T]\} \subset \mathbb{R}^m$, tenemos que considerar el siguiente problema de optimización:

$$(\hat{\mathbf{P}}_W^\ell) \quad \begin{aligned} & \min_{\tilde{\psi}_1, \dots, \tilde{\psi}_\ell \in \mathbb{R}^m} \int_0^T \left\| y(t) - \sum_{i=1}^{\ell} \langle y(t), \tilde{\psi}_i \rangle_W \tilde{\psi}_i \right\|_W^2 dt \\ & \text{s.t.} \quad \langle \tilde{\psi}_i, \tilde{\psi}_j \rangle_W = \delta_{ij} \text{ para } 1 \leq i, j \leq \ell. \end{aligned}$$

Vamos a resolver este problema usando las mismas técnicas que hemos empleado hasta ahora.

Observemos que para $\ell = 1$ tenemos el siguiente problema de mínimos:

$$\min_{\tilde{\psi} \in \mathbb{R}^m} \int_0^T \left\| y(t) - \langle y(t), \tilde{\psi} \rangle_W \tilde{\psi} \right\|_W^2 dt \quad \text{sujeito a} \quad \|\tilde{\psi}\|_W^2 = 1. \quad (2.28)$$

Supongamos que los $\{\tilde{\psi}_i\}_{i=2}^n$ son elegidos de tal forma que el conjunto $\{\tilde{\psi}, \tilde{\psi}_2, \dots, \tilde{\psi}_m\}$ sea una base ortonormal de \mathbb{R}^m con respecto al producto escalar $\langle \cdot, \cdot \rangle_W$. Entonces, tenemos:

$$y(t) = \langle y(t), \tilde{\psi} \rangle_W \tilde{\psi} + \sum_{i=2}^m \langle y(t), \tilde{\psi}_i \rangle_W \tilde{\psi}_i \quad \text{para todo } t \in [0, T].$$

Por tanto,

$$\begin{aligned} \int_0^T \left\| y(t) - \langle y(t), \tilde{\psi} \rangle_W \tilde{\psi} \right\|_W^2 dt &= \int_0^T \left\| \sum_{i=2}^m \langle y(t), \tilde{\psi}_i \rangle_W \tilde{\psi}_i \right\|_W^2 dt \\ &= \int_0^T \left\langle \sum_{i=2}^m \langle y(t), \tilde{\psi}_i \rangle_W \tilde{\psi}_i, \sum_{i=2}^m \langle y(t), \tilde{\psi}_i \rangle_W \tilde{\psi}_i \right\rangle_W dt \\ &= \sum_{i=2}^m \int_0^T |\langle y(t), \tilde{\psi}_i \rangle_W|^2 dt. \end{aligned}$$

Luego, (2.28) es equivalente al siguiente problema:

$$\min_{\tilde{\psi} \in \mathbb{R}^m} \int_0^T |\langle y(t), \tilde{\psi} \rangle_W|^2 dt \quad \text{sujeito a} \quad \|\tilde{\psi}\|_W^2 = 1. \quad (2.29)$$

Definimos la función Lagrangiana $\mathcal{L} : \mathbb{R}^m \times \mathbb{R} \rightarrow \mathbb{R}$ asociada a este problema, que viene dada por

$$\mathcal{L}(\psi, \lambda) = \int_0^T |\langle y(t), \psi \rangle_W|^2 dt + \lambda \left(1 - \|\psi\|_W^2 \right) \quad \text{para } (\psi, \lambda) \in \mathbb{R}^m \times \mathbb{R}.$$

Por los mismos razonamientos desarrollados en las Secciones 2.3 y 2.5 se puede probar que cualquier solución de (2.29) es un punto regular. Consecuentemente, las condiciones necesarias de optimalidad de primer orden viene dadas por

$$\nabla \mathcal{L}(\psi, \lambda) = 0 \quad \text{en } \mathbb{R}^m \times \mathbb{R}.$$

Por tanto, vamos a calcular la derivada parcial de \mathcal{L} con respecto a la i -ésima componente ψ_i del vector ψ :

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \psi_i}(\psi, \lambda) &= \frac{\partial}{\partial \psi_i} \left(\int_0^T \left| \sum_{k=1}^m \sum_{\nu=1}^m y_k(t) W_{k\nu} \psi_\nu \right|^2 dt + \lambda \left(1 - \sum_{k=1}^m \sum_{\nu=1}^m \psi_k W_{k\nu} \psi_\nu \right) \right) \\ &= 2 \int_0^T \left(\sum_{k=1}^m \sum_{\nu=1}^m y_k(t) W_{k\nu} \psi_\nu \right) \sum_{\mu=1}^m y_\mu(t) W_{\mu i} dt - 2\lambda \sum_{k=1}^m W_{ik} \psi_k \\ &= 2 \left(\int_0^T \langle y(t), \psi \rangle_W W y(t) dt - \lambda W \psi \right)_i \end{aligned}$$

para $i \in \{1, \dots, m\}$. Luego,

$$\nabla_\psi \mathcal{L}(\psi, \lambda) = 2 \left(\int_0^T \langle y(t), \psi \rangle_W W y(t) dt - \lambda W \psi \right).$$

Es decir, la condición $\nabla \mathcal{L}(\psi, \lambda) = 0$ nos conduce a la siguiente ecuación

$$\int_0^T \langle y(t), \psi \rangle_W W y(t) dt = \lambda W \psi \quad \text{en } \mathbb{R}^m. \quad (2.30)$$

Multiplicando (2.30) por W^{-1} por la izquierda obtenemos una expresión más simplificada:

$$\int_0^T \langle y(t), \psi \rangle_W y(t) dt = \lambda \psi \quad \text{en } \mathbb{R}^m. \quad (2.31)$$

A continuación, vamos a definir el operador $\mathcal{R} : \mathbb{R}^m \rightarrow \mathbb{R}^m$ como

$$\mathcal{R}(\psi) = \int_0^T \langle y(t), \psi \rangle_W y(t) dt, \quad \text{para } \psi \in \mathbb{R}^m$$

y vamos a obtener algunas de sus propiedades.

A partir de ahora nos referiremos a este operador como un operador continuo ya que a pesar de estar definido de \mathbb{R}^m en \mathbb{R}^m , viene dado a través de una integral.

Proposición 2.6.1. El operador \mathcal{R} es lineal y continuo. Además,

a) \mathcal{R} es no negativo, es decir,

$$\langle \mathcal{R}\psi, \psi \rangle_W \geq 0 \quad \text{para todo } \psi \in \mathbb{R}^m.$$

b) \mathcal{R} es autoadjunto, es decir,

$$\langle \mathcal{R}\psi, \tilde{\psi} \rangle_W = \langle \psi, \mathcal{R}\tilde{\psi} \rangle_W \quad \text{para todo } \psi, \tilde{\psi} \in \mathbb{R}^m.$$

Demostración. La linealidad del operador \mathcal{R} se da trivialmente por la linealidad del producto escalar y de la integral. Ahora bien, como \mathcal{R} es lineal, demostrar que \mathcal{R} es continuo es equivalente a probar que \mathcal{R} es acotado, es decir, existe una constante $C > 0$ tal que

$$\|\mathcal{R}\psi\|_W \leq C\|\psi\|_W, \quad \text{para todo } \psi \in \mathbb{R}^m. \quad (2.32)$$

En efecto, usando la desigualdad de Cauchy-Schwarz, deducimos que

$$\begin{aligned} \|\mathcal{R}\psi\|_W &\leq \int_0^T \|\langle y(t), \psi \rangle_W y(t)\|_W dt = \int_0^T |\langle y(t), \psi \rangle_W| \|y(t)\|_W dt \\ &\leq \int_0^T \|y(t)\|_W^2 \|\psi\|_W dt = \left(\int_0^T \|y(t)\|_W^2 dt \right) \|\psi\|_W = \|y\|_{L^2(0,T;\mathbb{R}^m)}^2 \|\psi\|_W \end{aligned}$$

para cualquier $\psi \in \mathbb{R}^m$. Ahora bien, como $y \in C([0, T]; \mathbb{R}^m) \subset L^2(0, T; \mathbb{R}^m)$, la norma $\|y\|_{L^2(0,T;\mathbb{R}^m)}$ está acotada y por ende se cumple (2.32) y \mathcal{R} está acotado.

Veamos ahora que \mathcal{R} es no negativo:

$$\begin{aligned} \langle \mathcal{R}\psi, \psi \rangle_W &= \left(\int_0^T \langle y(t), \psi \rangle_W y(t) dt \right)^\top W\psi = \int_0^T \langle y(t), \psi \rangle_W y(t)^\top W\psi dt \\ &= \int_0^T |\langle y(t), \psi \rangle_W|^2 dt \geq 0 \end{aligned}$$

para toda $\psi \in \mathbb{R}^m$.

Por último, veamos que \mathcal{R} es autoadjunto:

$$\begin{aligned} \langle \mathcal{R}\psi, \tilde{\psi} \rangle_W &= \int_0^T \langle y(t), \psi \rangle_W \langle y(t), \tilde{\psi} \rangle_W dt = \left\langle \int_0^T \langle y(t), \tilde{\psi} \rangle_W y(t) dt, \psi \right\rangle_W \\ &= \langle \mathcal{R}\tilde{\psi}, \psi \rangle_W = \langle \psi, \mathcal{R}\tilde{\psi} \rangle_W \end{aligned}$$

para toda $\psi, \tilde{\psi} \in \mathbb{R}^m$. □

Usando el operador \mathcal{R} , podemos reescribir (2.31) como el siguiente problema de autovalores:

$$\mathcal{R}\psi = \lambda\psi \quad \text{en } \mathbb{R}^m.$$

Además, por la Proposición 2.6.1, \mathcal{R} posee unos autovectores $\{\psi_i\}_{i=1}^m$ y autovalores reales asociados $\{\lambda_i\}_{i=1}^m$ tales que

$$\mathcal{R}\psi_i = \lambda_i\psi_i \quad \text{para } 1 \leq i \leq m \quad \text{y } \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m \geq 0. \quad (2.33)$$

Observemos que ψ_1 resuelve el problema (2.28), ya que

$$\begin{aligned} \int_0^T |\langle y(t), \psi_i \rangle_W|^2 dt &= \int_0^T \langle \langle y(t), \psi_i \rangle_W y(t), \psi_i \rangle_W dt = \langle \mathcal{R}\psi_i, \psi_i \rangle_W = \lambda_i \|\psi_i\|_W^2 \\ &= \lambda_i \quad \text{para } i \in \{1, \dots, m\} \end{aligned}$$

y λ_1 es el mayor de los autovalores.

El siguiente resultado, cuya prueba se basa en las mismas ideas que las ya desarrolladas en secciones anteriores nos muestra que la base POD de rango ℓ que resuelve el problema $(\hat{\mathbf{P}}_W^\ell)$ es la formada por los autovectores de \mathcal{R} correspondientes a los ℓ mayores autovalores.

Teorema 2.6.2. Supongamos que (SL) tiene una única solución $y : [0, T] \rightarrow \mathbb{R}^m$. Entonces, la base POD de rango ℓ que resuelve el problema de mínimos $(\hat{\mathbf{P}}_W^\ell)$ viene dada a través de los autovectores de \mathcal{R} correspondientes a los ℓ mayores autovalores $\lambda_1 \geq \dots \geq \lambda_\ell$.

Volvamos ahora a las condiciones de optimalidad dadas en (2.27), es decir,

$$YDY^\top W\psi_i = \lambda_i\psi_i \quad \text{para } i = 1, \dots, \ell.$$

Para cualquier $\psi \in \mathbb{R}^m$ e $i \in \{1, \dots, m\}$ obtenemos:

$$\begin{aligned} (YDY^\top W\psi)_i &= \sum_{\nu=1}^m \sum_{j=1}^m \sum_{k=1}^m \alpha_j Y_{ij} Y_{kj} W_{k\nu} \psi_\nu = \sum_{j=1}^n \alpha_j Y_{ij} \langle y_j, \psi \rangle_W \\ &= \sum_{j=1}^n \alpha_j \langle y_j, \psi \rangle_W (y_j)_i \end{aligned}$$

donde $(y_j)_i$ denota la i -ésima componente del vector $y_j \in \mathbb{R}^m$. Por tanto,

$$YDY^\top W\psi = \sum_{j=1}^n \alpha_j \langle y_j, \psi \rangle_W y_j.$$

Definimos ahora el operador $\mathcal{R}^n : \mathbb{R}^m \rightarrow \mathbb{R}^m$ como

$$\mathcal{R}^n(\psi) = \sum_{j=1}^n \alpha_j \langle y_j, \psi \rangle_W y_j \quad \text{para } \psi \in \mathbb{R}^m$$

y vamos a obtener algunas de sus propiedades.

Observemos, que a diferencia de \mathcal{R} , este operador es discreto porque está definido a partir de un suma.

Proposición 2.6.3. El operador \mathcal{R}^n es lineal y continuo. Además,

a) \mathcal{R}^n es no negativo, es decir,

$$\langle \mathcal{R}^n \psi, \psi \rangle_W \geq 0 \quad \text{para todo } \psi \in \mathbb{R}^m.$$

b) \mathcal{R}^n es autoadjunto, es decir,

$$\langle \mathcal{R}^n \psi, \tilde{\psi} \rangle_W = \langle \psi, \mathcal{R}^n \tilde{\psi} \rangle_W \quad \text{para todo } \psi, \tilde{\psi} \in \mathbb{R}^m.$$

Demostración. La prueba de este resultado sigue el mismo razonamiento que el ya efectuado en la Proposición 2.6.1.

La linealidad del operador \mathcal{R}^n se tiene trivialmente por la linealidad del producto escalar. Como acabos de probar que \mathcal{R}^n es lineal, demostrar que \mathcal{R}^n es continuo es equivalente a probar que \mathcal{R}^n es acotado, es decir, existe una constante $C > 0$ tal que

$$\|\mathcal{R}^n \psi\|_W \leq C \|\psi\|_W, \quad \text{para todo } \psi \in \mathbb{R}^m. \quad (2.34)$$

En efecto, usando la desigualdad de Cauchy-Schwarz, deducimos que

$$\begin{aligned} \|\mathcal{R}^n \psi\|_W &\leq \sum_{j=1}^n \|\alpha_j \langle y_j, \psi \rangle_W y_j\| \leq \sum_{j=1}^n \alpha_j |\langle y_j, \psi \rangle_W| \|y_j\|_W \leq \sum_{j=1}^n \alpha_j \|y_j\|_W^2 \|\psi\|_W \\ &= \|\psi\|_W \sum_{j=1}^n \alpha_j \|y_j\|_W^2 = C \|\psi\|_W \end{aligned}$$

para cualquier $\psi \in \mathbb{R}^m$. Es decir, \mathcal{R}^n está acotado.

Veamos ahora que \mathcal{R}^n es no negativo:

$$\langle \mathcal{R}^n \psi, \psi \rangle_W = \left\langle \sum_{j=1}^n \alpha_j \langle y_j, \psi \rangle_W y_j, \psi \right\rangle_W = \sum_{j=1}^n \alpha_j |\langle y_j, \psi \rangle_W|^2 \geq 0$$

para todo $\psi \in \mathbb{R}^m$.

Por último, probemos que es autoadjunto:

$$\begin{aligned} \langle \mathcal{R}^n \psi, \tilde{\psi} \rangle_W &= \left\langle \sum_{j=1}^n \alpha_j \langle y_j, \psi \rangle_W y_j, \tilde{\psi} \right\rangle_W = \sum_{j=1}^n \alpha_j \langle y_j, \psi \rangle_W \langle y_j, \tilde{\psi} \rangle_W \\ &= \left\langle \sum_{j=1}^n \alpha_j \langle y_j, \tilde{\psi} \rangle_W y_j, \psi \right\rangle_W = \langle \mathcal{R}^n \tilde{\psi}, \psi \rangle_W = \langle \psi, \mathcal{R}^n \tilde{\psi} \rangle_W \end{aligned}$$

para todo $\psi, \tilde{\psi} \in \mathbb{R}^m$. □

Por tanto, \mathcal{R}^n tiene las mismas propiedades que el operador \mathcal{R} anteriormente definido.

Resumiendo, tenemos:

$$\mathcal{R}^n \psi_i^n = \lambda_i^n \psi_i^n, \quad \lambda_1^n \geq \dots \geq \lambda_\ell^n \geq \dots \geq \lambda_{d(n)}^n > \lambda_{d(n)+1}^n = \dots = \lambda_m^n = 0, \quad (2.35)$$

$$\mathcal{R} \psi_i = \lambda_i \psi_i, \quad \lambda_1 \geq \dots \geq \lambda_\ell \geq \dots \geq \lambda_d > \lambda_{d+1} = \dots = \lambda_m = 0. \quad (2.36)$$

Lema 2.6.4. Bajo las condiciones anteriores, se cumplen las siguientes dos igualdades:

a)

$$\int_0^T \|y(t)\|_W^2 dt = \sum_{i=1}^d \lambda_i = \sum_{i=1}^m \lambda_i.$$

b)

$$\sum_{j=1}^n \alpha_j \|y(t_j)\|_W^2 = \sum_{j=1}^{d(n)} \lambda_j^n = \sum_{j=1}^m \lambda_j^n, \quad \text{para cada } n \in \mathbb{N}.$$

Demostración. Vamos a comenzar probando el apartado a). Recordemos que

$$\mathcal{R} \psi_i = \int_0^T \langle y(t), \psi_i \rangle_W y(t) dt \quad \text{para todo } i \in \{1, \dots, m\}.$$

Si tomamos producto escalar, con respecto a W con ψ_i , obtenemos:

$$\begin{aligned} \langle \mathcal{R}\psi_i, \psi_i \rangle_W &= \left(\int_0^T \langle y(t), \psi_i \rangle_W y(t) dt \right)^\top W \psi_i = \int_0^T \langle y(t), \psi_i \rangle_W y(t)^\top W \psi_i dt \\ &= \int_0^T |\langle y(t), \psi_i \rangle_W|^2 dt. \end{aligned}$$

Ahora, sumamos desde $i = 1, \dots, d$:

$$\sum_{i=1}^d \int_0^T |\langle y(t), \psi_i \rangle_W|^2 dt = \sum_{i=1}^d \langle \mathcal{R}\psi_i, \psi_i \rangle_W = \sum_{i=1}^d \langle \lambda_i \psi_i, \psi_i \rangle_W = \sum_{i=1}^d \lambda_i = \sum_{i=1}^m \lambda_i.$$

Si desarrollamos $y(t) \in \mathbb{R}^m$ en términos de $\{\psi_i\}_{i=1}^m$, tenemos

$$y(t) = \sum_{i=1}^m \langle y(t), \psi_i \rangle_W \psi_i$$

y por tanto,

$$\int_0^T \|y(t)\|_W^2 dt = \sum_{i=1}^m \int_0^T |\langle y(t), \psi_i \rangle_W|^2 dt = \sum_{i=1}^m \lambda_i$$

Veamos, por último, que se cumple b). En efecto,

$$\mathcal{R}^n \psi_i^n = \sum_{j=1}^n \alpha_j \langle y_j, \psi_i^n \rangle_W y_j \quad \text{para todo } i \in \{1, \dots, m\} \text{ y } n \in \mathbb{N}.$$

Multiplicamos escalarmente con respecto a W por ψ_i^n :

$$\langle \mathcal{R}^n \psi_i^n, \psi_i^n \rangle_W = \left\langle \sum_{j=1}^n \alpha_j \langle y_j, \psi_i^n \rangle_W y_j, \psi_i^n \right\rangle_W = \sum_{j=1}^n \alpha_j |\langle y_j, \psi_i^n \rangle_W|^2.$$

Ahora, sumamos desde $i = 1, \dots, d(n)$:

$$\begin{aligned} \sum_{i=1}^{d(n)} \sum_{j=1}^n \alpha_j |\langle y_j, \psi_i^n \rangle_W|^2 &= \sum_{i=1}^{d(n)} \langle \mathcal{R}^n \psi_i^n, \psi_i^n \rangle_W \\ &= \sum_{i=1}^{d(n)} \langle \lambda_i^n \psi_i^n, \psi_i^n \rangle_W \\ &= \sum_{i=1}^{d(n)} \lambda_i^n = \sum_{i=1}^m \lambda_i^n. \end{aligned}$$

Si desarrollamos $y_j \in \mathbb{R}^m$ en términos de $\{\psi_i^n\}_{i=1}^m$, tenemos

$$y_j = \sum_{i=1}^m \langle y_j, \psi_i^n \rangle_W \psi_i^n,$$

y por tanto,

$$\sum_{j=1}^n \alpha_j \|y(t_j)\|_W^2 = \sum_{i=1}^m \sum_{j=1}^n \alpha_j |\langle y_j, \psi_i^n \rangle_W|^2 = \sum_{i=1}^m \lambda_i^n \quad \text{para todo } n \in \mathbb{N}.$$

□

Llegados a este punto nos proponemos demostrar que el operador discreto \mathcal{R}^n converge al operador continuo \mathcal{R} y que las correspondientes sucesiones de autovalores y autovectores también lo hacen.

Teorema 2.6.5. Supongamos que (SL) tiene una única solución $y : [0, T] \rightarrow \mathbb{R}^m$, $y \in \mathcal{C}^1(0, T; \mathbb{R}^m) \cap \mathcal{C}([0, T]; \mathbb{R}^m)$ y sean $\{(\psi_i^n, \lambda_i^n)\}_{i=1}^m$ y $\{(\psi_i, \lambda_i)\}_{i=1}^m$ los pares autovectores-autovalores dados en (2.35) y (2.36). Supongamos que $\ell \in \{1, \dots, m\}$ es un valor fijado verificando $\lambda_\ell \neq \lambda_{\ell+1}$ y

$$\sum_{i=\ell+1}^m \lambda_i \neq 0, \quad \sum_{i=\ell+1}^m |\langle y_0, \psi_i \rangle_W|^2 \neq 0.$$

Entonces, se cumple que

$$\lim_{n \rightarrow \infty} \|\mathcal{R}^n - \mathcal{R}\|_{\mathcal{L}(\mathbb{R}^m)} = 0, \quad (2.37)$$

donde $\mathcal{L}(\mathbb{R}^m)$ denota el espacio de Banach de todas las aplicaciones lineales y continuas de \mathbb{R}^m en sí mismo y en consecuencia:

$$\begin{aligned} \lim_{n \rightarrow \infty} |\lambda_i^n - \lambda_i| &= \lim_{n \rightarrow \infty} \|\psi_i^n - \psi_i\|_W = 0 \quad \text{para } 1 \leq i \leq \ell, \\ \lim_{n \rightarrow \infty} \sum_{i=\ell+1}^m (\lambda_i^n - \lambda_i) &= 0, \\ \lim_{n \rightarrow \infty} \sum_{i=\ell+1}^m |\langle y_0, \psi_i^n \rangle_W|^2 &= \sum_{i=\ell+1}^m |\langle y_0, \psi_i \rangle_W|^2. \end{aligned}$$

Demostración. Sea $\psi \in \mathbb{R}^m$ un vector arbitrario tal que $\|\psi\|_W = 1$. Definamos la función $f_\psi : [0, T] \rightarrow \mathbb{R}^m$ por:

$$f_\psi(t) = \langle y(t), \psi \rangle_W y(t) \quad \text{para } t \in [0, T].$$

Entonces, tenemos que $f_\psi \in C^1([0, T]; \mathbb{R}^m)$ con

$$\dot{f}_\psi(t) = \langle \dot{y}(t), \psi \rangle_W y(t) + \langle y(t), \psi \rangle_W \dot{y}(t) \quad \text{para } t \in [0, T].$$

Usando el desarrollo de Taylor para la función $f_\psi(t)$, existen $\tau_{j1}(t) \in (t_j, t)$ y $\tau_{j2}(t) \in (t, t_{j+1})$ dependientes de t tales que:

$$f_\psi(t) = f_\psi(t_j) + \dot{f}_\psi(\tau_{j1}(t))(t - t_j) \quad \text{y} \quad f_\psi(t) = f_\psi(t_{j+1}) + \dot{f}_\psi(\tau_{j2}(t))(t - t_{j+1}).$$

Entonces:

$$\begin{aligned} \int_{t_j}^{t_{j+1}} f_\psi(t) dt &= \frac{1}{2} \int_{t_j}^{t_{j+1}} f_\psi(t_j) + \dot{f}_\psi(\tau_{j1}(t))(t - t_j) dt \\ &\quad + \frac{1}{2} \int_{t_j}^{t_{j+1}} f_\psi(t_{j+1}) + \dot{f}_\psi(\tau_{j2}(t))(t - t_{j+1}) dt \\ &= \frac{\Delta t}{2} (f_\psi(t_j) + f_\psi(t_{j+1})) + \frac{1}{2} \int_{t_j}^{t_{j+1}} \dot{f}_\psi(\tau_{j1}(t))(t - t_j) dt \\ &\quad + \frac{1}{2} \int_{t_j}^{t_{j+1}} \dot{f}_\psi(\tau_{j2}(t))(t - t_{j+1}) dt \end{aligned}$$

Por tanto,

$$\begin{aligned} \|\mathcal{R}^n \psi - \mathcal{R} \psi\|_W &= \left\| \sum_{j=1}^n \alpha_j f_\psi(t_j) - \int_0^T f_\psi(t) dt \right\|_W \\ &= \left\| \sum_{j=1}^{n-1} \left(\frac{\Delta t}{2} (f_\psi(t_j) + f_\psi(t_{j+1})) - \int_{t_j}^{t_{j+1}} f_\psi(t) dt \right) \right\|_W \\ &\leq \frac{1}{2} \sum_{j=1}^{n-1} \int_{t_j}^{t_{j+1}} \|\dot{f}_\psi(\tau_{j1}(t))\|_W |t - t_j| + \|\dot{f}_\psi(\tau_{j2}(t))\|_W |t - t_{j+1}| dt \\ &\leq \frac{1}{2} \max_{t \in [0, T]} \|\dot{f}_\psi(t)\|_W \sum_{j=1}^{n-1} \left(\frac{(t - t_j)^2}{2} - \frac{(t_{j+1} - t)^2}{2} \right) \Bigg|_{t=t_j}^{t=t_{j+1}} \\ &= \frac{1}{2} \max_{t \in [0, T]} \|\dot{f}_\psi(t)\|_W \sum_{j=1}^{n-1} (\Delta t)^2 = \frac{(\Delta t)^2 (n-1)}{2} \max_{t \in [0, T]} \|\dot{f}_\psi(t)\|_W \\ &= \frac{\Delta t T}{2} \max_{t \in [0, T]} \|\dot{f}_\psi(t)\|_W \\ &= \frac{\Delta t T}{2} \max_{t \in [0, T]} \|\langle \dot{y}(t), \psi \rangle_W y(t) + \langle y(t), \psi \rangle_W \dot{y}(t)\|_W \\ &= \Delta t T \max_{t \in [0, T]} \|\dot{y}(t)\|_W \|y(t)\|_W \leq \Delta t T \|y\|_{C^1([0, T]; \mathbb{R}^m)}^2. \end{aligned}$$

Luego,

$$\|\mathcal{R}^n - \mathcal{R}\|_{\mathcal{L}(\mathbb{R}^m)} = \sup_{\|\psi\|_W=1} \|\mathcal{R}^n \psi - \mathcal{R} \psi\|_W \leq \Delta t T \|y\|_{C^1([0,T];\mathbb{R}^m)}^2 \xrightarrow{\Delta t \rightarrow 0} 0.$$

Vamos a probar a continuación que

$$\sum_{i=1}^m \lambda_i^n \rightarrow \sum_{i=1}^m \lambda_i \text{ cuando } n \rightarrow \infty. \quad (2.38)$$

En efecto, por el Lema 2.6.4, la convergencia dada en (2.38) se puede reescribir como

$$\sum_{j=1}^n \alpha_j \|y(t_j)\|_W^2 \rightarrow \int_0^T \|y(t)\|_W^2 dt \text{ cuando } \Delta t \rightarrow 0. \quad (2.39)$$

Luego, basta elegir los α_j de manera que se garantice dicha convergencia. Para ello, si aproximamos esa integral usando la regla del trapecio,

$$\int_0^T \|y(t)\|_W^2 dt = \sum_{j=1}^{n-1} \int_{t_j}^{t_{j+1}} \|y(t)\|_W^2 dt \approx \sum_{j=1}^{n-1} \frac{\Delta t}{2} \left(\|y(t_j)\|_W^2 + \|y(t_{j+1})\|_W^2 \right),$$

los pesos α_j ($j = 1, \dots, n$) que tendríamos que poner para tener la convergencia buscada serían los siguientes:

$$\alpha_1 = \frac{\Delta t}{2}, \alpha_j = \Delta t \text{ para } 2 \leq j \leq n-1, \alpha_n = \frac{\Delta t}{2}. \quad (2.40)$$

Ahora, fijemos $\ell \in \{1, \dots, m\}$ tal que $\lambda_\ell \neq \lambda_{\ell+1}$. Entonces, por el análisis espectral de operadores compactos y (2.37), se sigue que

$$\lambda_i^n \rightarrow \lambda_i \text{ para } 1 \leq i \leq \ell \text{ cuando } n \rightarrow \infty. \quad (2.41)$$

Combinando (2.38) y (2.41), existe un $n_0 \in \mathbb{N}$ tal que

$$\sum_{i=\ell+1}^m \lambda_i^n \leq 2 \sum_{i=\ell+1}^m \lambda_i \text{ para todo } n \geq n_0 \quad (2.42)$$

si $\sum_{i=\ell+1}^m \lambda_i \neq 0$. Además, para ℓ como antes, n_0 puede ser escogido de forma que se verifique:

$$\sum_{i=\ell+1}^{d(n)} |\langle y_0, \psi_i^n \rangle_W|^2 \leq 2 \sum_{i=\ell+1}^m |\langle y_0, \psi_i \rangle_W|^2 \text{ para todo } n \geq n_0 \quad (2.43)$$

siempre que $\sum_{i=\ell+1}^m |\langle y_0, \psi_i \rangle_W|^2 \neq 0$ y se cumpla (2.37).

En efecto, recordemos que el vector $y_0 \in \mathbb{R}^m$ es la condición inicial dada en el problema (SL). Por tanto, se verifica que

$$\|y_0\|_W^2 = \sum_{i=1}^m |\langle y_0, \psi_i \rangle_W|^2. \quad (2.44)$$

Si $t_1 = 0$, tenemos que $y_0 \in \langle y_1, \dots, y_n \rangle$ para cada n y

$$\|y_0\|_W^2 = \sum_{i=1}^{d(n)} |\langle y_0, \psi_i^n \rangle_W|^2. \quad (2.45)$$

Luego, para $\ell < d(n)$ y usando (2.44) y (2.45), es fácil comprobar que

$$\begin{aligned} \sum_{i=\ell+1}^{d(n)} |\langle y_0, \psi_i^n \rangle_W|^2 &= \sum_{i=1}^{d(n)} |\langle y_0, \psi_i^n \rangle_W|^2 - \sum_{i=1}^{\ell} |\langle y_0, \psi_i^n \rangle_W|^2 + \sum_{i=1}^{\ell} |\langle y_0, \psi_i \rangle_W|^2 \\ &\quad + \sum_{i=\ell+1}^m |\langle y_0, \psi_i \rangle_W|^2 - \sum_{i=1}^m |\langle y_0, \psi_i \rangle_W|^2 \\ &= \sum_{i=1}^{\ell} \left(|\langle y_0, \psi_i \rangle_W|^2 - |\langle y_0, \psi_i^n \rangle_W|^2 \right) + \sum_{i=\ell+1}^m |\langle y_0, \psi_i \rangle_W|^2. \end{aligned}$$

Como consecuencia de (2.37) y de que ℓ es tal que $\lambda_\ell \neq \lambda_{\ell+1}$, tenemos que

$\lim_{n \rightarrow \infty} \|\psi_i^n - \psi_i\|_W = 0$ para $i = 1, \dots, \ell$ y por ende se verifica (2.43). \square

Nota 2.6.6. Hemos realizado el estudio teórico sobre el problema (SL), pero podríamos haberlo hecho para cualquier otro problema realizando argumentos análogos a los ya descritos.

Cálculo efectivo de la base POD

Definamos el operador lineal y continuo $\mathcal{Y} : L^2(0, T) \rightarrow \mathbb{R}^m$ de la siguiente forma.

$$\mathcal{Y}\phi := \int_0^T \phi(t)y(t) dt \quad \text{para } \phi \in L^2(0, T).$$

Observemos que \mathcal{Y} es un operador compacto ya que es de rango finito (Teorema 1.2.10).

Definamos ahora su operador adjunto $\mathcal{Y}^* : \mathbb{R}^m \rightarrow L^2(0, T)$ como

$$\langle \mathcal{Y}^*\psi, \phi \rangle_{L^2(0, T)} = \langle \psi, \mathcal{Y}\phi \rangle_W \quad \text{para todo } (\psi, \phi) \in \mathbb{R}^m \times L^2(0, T).$$

Usando esta definición, es fácil comprobar que $(\mathcal{Y}^*\psi)(t) = \langle \psi, y(t) \rangle_W$ para todo $\psi \in \mathbb{R}^m$ y para casi todo $t \in [0, T]$. Además, \mathcal{Y}^* es compacto al serlo \mathcal{Y} (Teorema 1.2.7).

Entonces, tenemos:

$$\mathcal{Y}\mathcal{Y}^*\psi = \int_0^T \langle \psi, y(t) \rangle_W y(t) dt = \int_0^T \langle y(t), \psi \rangle_W y(t) dt = \mathcal{R}\psi$$

para todo $\psi \in \mathbb{R}^m$; es decir, $\mathcal{R} = \mathcal{Y}\mathcal{Y}^*$. Más aún,

$$(\mathcal{Y}^*\mathcal{Y}\phi)(t) = \left\langle \int_0^T \phi(s)y(s) ds, y(t) \right\rangle_W = \int_0^T \langle y(s), y(t) \rangle_W \phi(s) ds =: (\mathcal{K}\phi)(t)$$

para todo $\phi \in L^2(0, T)$ y en casi todo $t \in [0, T]$. Por tanto, $\mathcal{K} = \mathcal{Y}^*\mathcal{Y}$.

Se puede probar que el operador $\mathcal{K} : L^2(0, T) \rightarrow L^2(0, T)$ es lineal, acotado, auto-adjunto y compacto (esto último es trivial por el Teorema 1.2.9). Por tanto, la base POD puede también ser computada de la siguiente forma: Resolver

$$\mathcal{K}\phi_i = \lambda_i\phi_i \text{ para } 1 \leq i \leq \ell, \quad \lambda_1 \geq \dots \geq \lambda_\ell > 0, \quad \int_0^T \phi_i(t)\phi_j(t) dt = \delta_{ij} \quad (2.46)$$

y tomar

$$\psi_i = \frac{1}{\sqrt{\lambda_i}}\mathcal{Y}\phi_i = \frac{1}{\sqrt{\lambda_i}} \int_0^T \phi_i(t)y(t) dt \quad \text{para } i = 1, \dots, \ell.$$

Observemos que (2.46) es un problema de autovalores en el espacio de dimensión infinita $L^2(0, T)$.

2.6.2. Ecuaciones en derivadas parciales

Vamos a considerar la ecuación del calor uno-dimensional:

$$(EC) \quad \begin{cases} \theta_t(t, x) = \theta_{xx}(t, x) & \text{para todo } (t, x) \in (0, T) \times \Omega, \\ \theta_x(t, 0) = \theta_x(t, 1) = 0 & \text{para todo } t \in (0, T), \\ \theta(0, x) = \theta_0(x) & \text{para todo } x \in \Omega, \end{cases}$$

donde $\Omega = (0, 1)$, $T > 0$ y $\theta_0 \in \mathcal{C}(\bar{\Omega})$ es una condición inicial dada. Para resolver (EC) numéricamente vamos a aplicar una aproximación por diferencias finitas en la variable espacial x . Igual que hicimos en el Ejemplo 2.5.1, vamos a construir una partición uniforme $\{x_i\}_{i=1}^m$ del intervalo $[0, 1]$. Denotemos por $y_i : [0, T] \rightarrow \mathbb{R}$ a la

aproximación numérica de $\theta(\cdot, x_i)$ para $i = 1, \dots, m$. La derivada θ_{xx} y la condición de contorno son discretizadas usando diferencias finitas centradas de segundo orden; así que obtenemos las siguientes ecuaciones diferenciales ordinarias para las funciones y_i :

$$\begin{cases} \dot{y}_1(t) = \frac{-2y_1(t) + 2y_2(t)}{h^2}, \\ \dot{y}_i(t) = \frac{y_{i-1}(t) - 2y_i(t) + y_{i+1}(t)}{h^2}, & i = 2, \dots, m-1, \\ \dot{y}_m(t) = \frac{-2y_m(t) + 2y_{m-1}(t)}{h^2}, \end{cases}$$

para $t \in (0, T]$. Además, de $\theta(0, x) = \theta_0(x)$, inferimos la condición inicial $y_i(0) = \theta_0(x_i)$, $i = 1, \dots, m$. Vamos a expresar lo anterior como un sistema de ecuaciones diferenciales. Para ello, introducimos la matriz

$$A = \frac{1}{h^2} \begin{pmatrix} -2 & 2 & & & 0 \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & & 1 & -2 & 1 \\ 0 & & & & 2 & -2 \end{pmatrix} \in \mathbb{R}^{m \times m}$$

y los vectores

$$y(t) = \begin{pmatrix} y_1(t) \\ \vdots \\ y_m(t) \end{pmatrix} \text{ para } t \in [0, T], \quad y_0 = \begin{pmatrix} \theta_0(x_1) \\ \vdots \\ \theta_0(x_m) \end{pmatrix} \in \mathbb{R}^m,$$

de modo que podemos escribir:

$$\begin{cases} \dot{y}(t) = Ay(t) & \text{para } t \in (0, T], \\ y(0) = y_0. \end{cases}$$

Observemos que este problema lineal de valor inicial coincide con (SL) tomando $f = 0$. Además ahora el vector $y(t)$, $t \in [0, T]$, representa una función en Ω evaluada en m puntos. Por tanto, deberíamos dotar \mathbb{R}^m por un producto escalar con peso que represente un producto escalar discreto en un apropiado espacio de funciones. Aquí, elegimos el producto escalar introducido en (2.23). Después elegimos una partición

$\{t_j\}_{j=1}^n$ del intervalo $[0, T]$ y definimos $y_j = y(t_j)$ para $j = 1, \dots, n$. Si estamos interesados en encontrar una base POD de rango $\ell \leq \min\{m, n\}$ que describa el conjunto de puntos $\{y_j\}_{j=1}^n$, debemos plantear el problema $(\hat{\mathbf{P}}_W^{n, \ell})$.

Capítulo 3

Aplicaciones prácticas del método POD

En este capítulo vamos a aplicar los resultados teóricos que hemos tratado con anterioridad en algunos casos prácticos, considerando un ejemplo para el caso discreto y otro para el caso continuo. Para ello, usaremos el software matemático MATLAB.

3.1. Caso discreto

Vamos a considerar una serie de patios de los que tenemos mediciones experimentales de sus temperaturas internas cada hora en un determinado período del año. Más concretamente, para que las temperaturas se mueven en un rango similar, hemos escogido patios en una misma zona climática cuyas mediciones se corresponden a períodos similares del año (en concreto, verano - otoño). A continuación, enumeramos los patios escogidos y las mediciones de las semanas consideradas.

1. Carlos Rubio en Córdoba (26 de julio a 1 de agosto).
2. Carlos Rubio en Córdoba (9-15 de agosto).
3. Hernán Cortés en Sevilla (5-11 de septiembre).
4. Hernán Cortés² en Sevilla (5-11 de septiembre).
5. IMUS, Instituto de Matemáticas de la Universidad de Sevilla (14-20 de agosto).
6. IMUS, Instituto de Matemáticas de la Universidad de Sevilla (4-10 de septiembre).

7. UNED, Universidad Nacional de Educación a Distancia en Sevilla (14-20 de agosto).
8. UNED, Universidad Nacional de Educación a Distancia en Sevilla (21-27 de agosto).

Nuestro objetivo, por tanto, será calcular una base POD para esta serie de datos experimentales que nos permita aproximar las temperaturas de un nuevo patio en un período del año similar.

Para ello, vamos a guiarnos del cálculo efectivo de la base POD descrito en la página 26; concretamente usaremos (2.18) y (2.19). Es decir, vamos a construir la matriz $Y = (y_1 | \dots | y_8) \in \mathbb{R}^{168 \times 8}$, donde cada columna y_j para $j = 1, \dots, 8$ representa las temperaturas de un patio medidas cada hora en la semana correspondiente, es decir, es un vector con 168 componentes. Posteriormente, calculamos los autovalores y autovectores de la matriz $Y^T Y$ y calculamos el valor de ℓ observando la función $\varepsilon(\ell)$ definida en (2.21).

En la siguiente tabla hemos recogido los valores de la función $\varepsilon(\ell)$ para cada $\ell = 1, \dots, 8$.

$\varepsilon(1)$	$\varepsilon(2)$	$\varepsilon(3)$	$\varepsilon(4)$	$\varepsilon(5)$	$\varepsilon(6)$	$\varepsilon(7)$	$\varepsilon(8)$
99.689 %	99.863 %	99.930 %	99.972 %	99.990 %	99.995 %	99.998 %	100 %

Tabla 1: Valores de la función $\varepsilon(\ell)$.

Como podemos observar todos estos valores están muy próximos al 100 %, lo que quiere decir que con muy pocas bases podemos capturar casi la totalidad de la energía del sistema. Por ello, vamos a elegir, por ejemplo, $\ell = 2$.

Para terminar, vamos a considerar un nuevo patio (IMUS, 21-27 de agosto) y observemos qué aproximación sobre las temperaturas internas obtenemos gracias a la base POD. Si y es el vector de las temperaturas internas de este nuevo patio, lo podemos aproximar como combinación lineal de los vectores de la base POD de la forma:

$$y \approx \langle y_1, \psi_1 \rangle \psi_1 + \langle y_2, \psi_2 \rangle \psi_2,$$

donde ψ_1 y ψ_2 son las funciones de base definidas en (2.19) y $\langle \cdot, \cdot \rangle$ denota el producto escalar euclídeo entre vectores.

Podemos comparar esta aproximación con el siguiente gráfico que muestra las temperaturas internas del patio nuevo con la aproximación obtenida con el método POD:

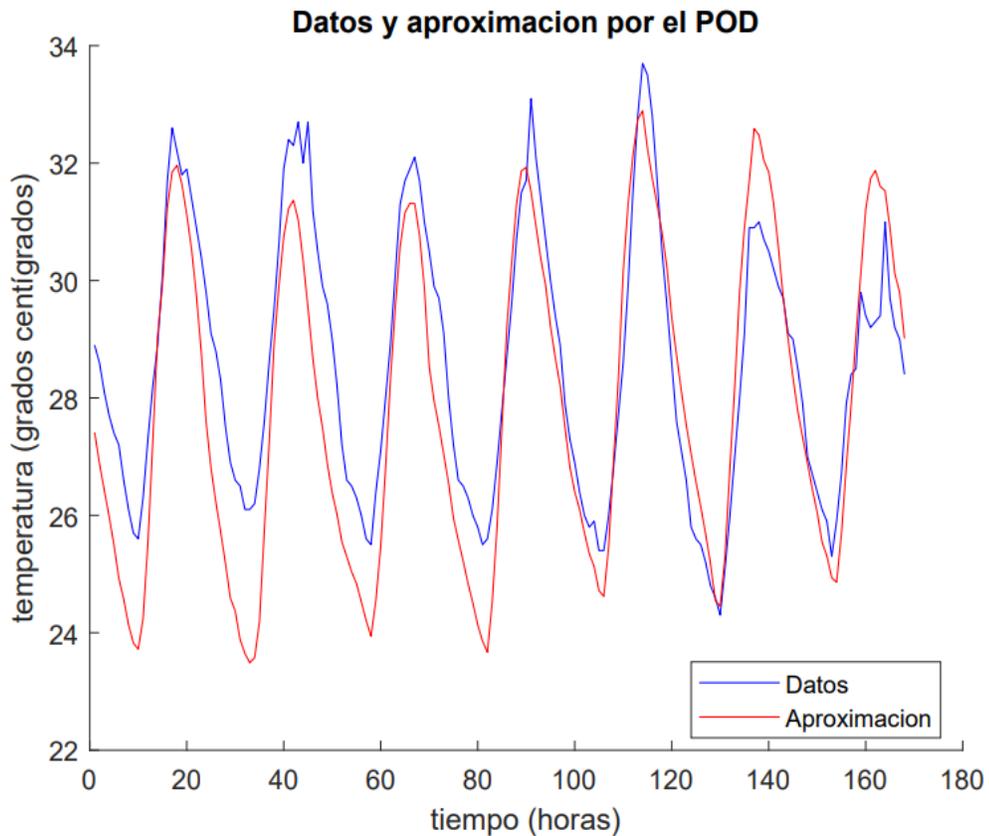


Figura 3.1: Temperaturas internas de los datos y de la aproximación por el POD.

Como podemos observar, con tan solo dos elementos en la base POD, obtenemos una buena aproximación del nuevo patio en el sentido de que el error cometido es pequeño. Todos estos datos experimentales han sido suministrados por el grupo de investigación Sostenibilidad en Arquitectura, Tecnología y Patrimonio: Materialidad y Sistemas Constructivos de la Universidad de Sevilla.

3.2. Caso continuo

Vamos a considerar el siguiente sistema diferencial ordinario de tres ecuaciones con tres incógnitas en el intervalo $[0, 30]$:

$$\begin{cases} y_1'(t) = y_2(t) y_3(t), \\ y_2'(t) = -0,7y_1(t) y_3(t), \\ y_3'(t) = -0,51y_1(t) y_2(t), \\ y_1(0) = 0; y_2(0) = y_3(0) = 1. \end{cases}$$

Nuestro objetivo es comparar la solución obtenida resolviendo el sistema numéricamente por un método de Runge-Kutta con la obtenida por el método POD; así como calcular aproximaciones en ciertos instantes de tiempo.

Consideremos una partición uniforme del intervalo $[0, 30]$ formada por 100 nodos, es decir, $\{0 = t_1 < t_2 < \dots < t_{100} = 30\}$. En la Figura 3.2 se muestra la solución aproximada del problema usando el método de Runge-Kutta de cuarto orden (ode45 de MATLAB).

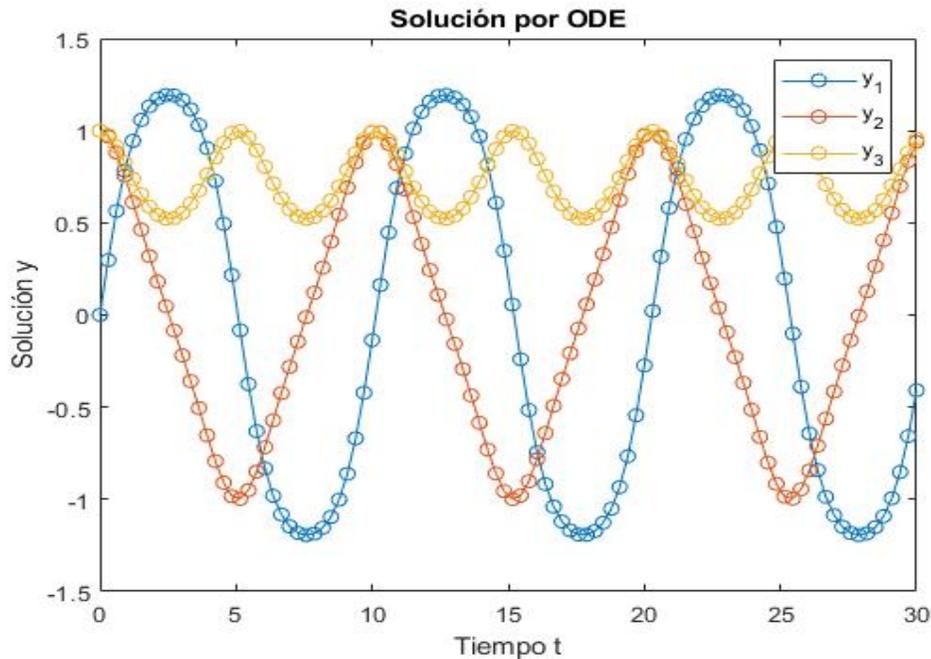


Figura 3.2: Solución aproximada usando ode45.

Ahora, vamos a realizar una aproximación numérica usando el método POD. Para ello, vamos a considerar una matriz Y formada por las tres componentes de la solución (y_1, y_2, y_3) definidas en cada uno de los 100 nodos de la partición del intervalo $[0, 30]$. Es decir, en la fila i -ésima ($i = 1, 2, 3$) de dicha matriz aparece $y_i(t_j)$ para cada $j = 1, \dots, 100$. A partir de esta matriz, calculamos los autovalores y autovectores de $Y^T Y$ y estudiamos los valores de la función $\varepsilon(\ell)$ considerando una tolerancia del 80%. En las Figuras 3.3, 3.4 y 3.5 podemos visualizar la aproximación de cada componente de la solución usando N funciones de base. Se observa que tomando tres funciones de base, la aproximación obtenida es bastante buena, en sentido de que el error que estamos cometiendo es muy pequeño.

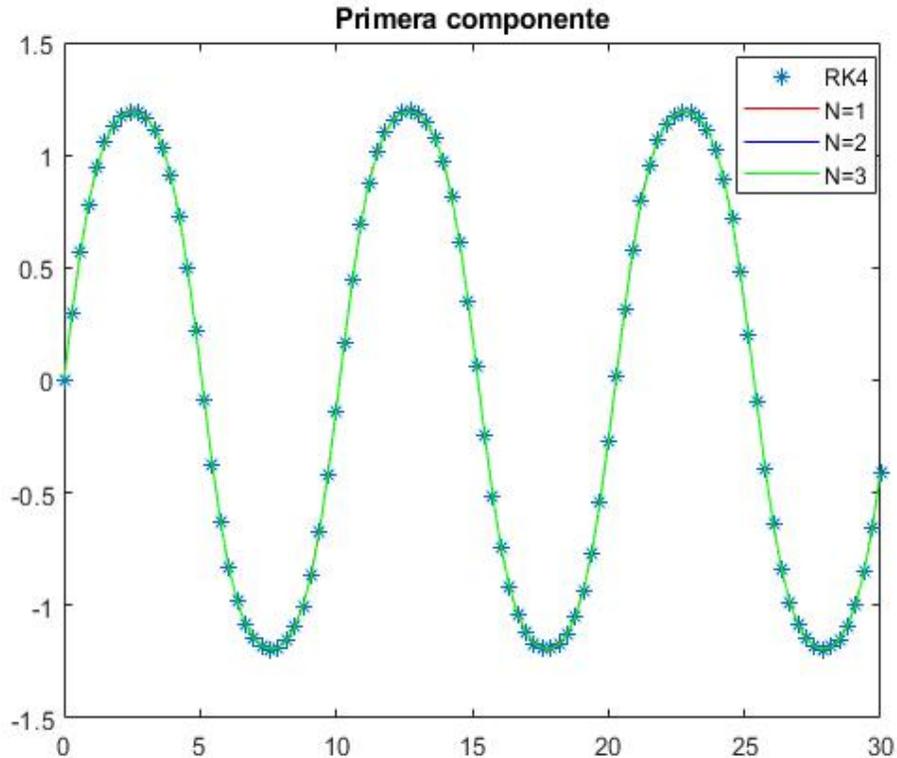


Figura 3.3: Aproximación de la primera componente y_1 .

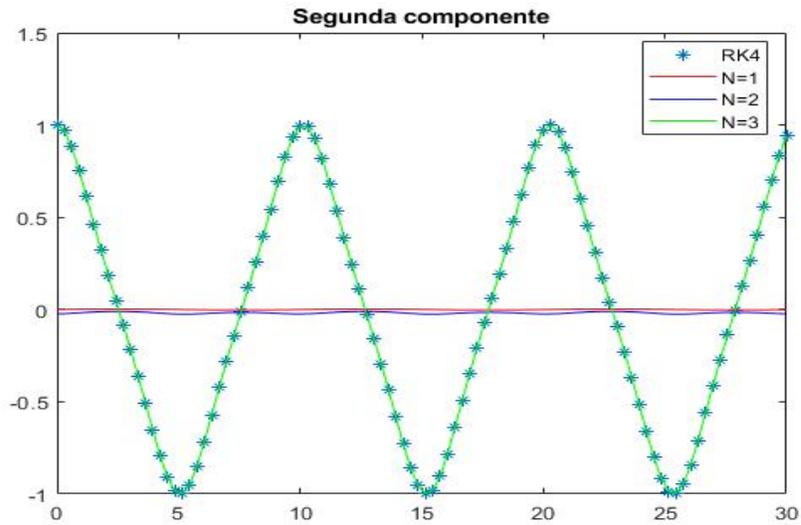


Figura 3.4: Aproximación de la segunda componente y_2 .

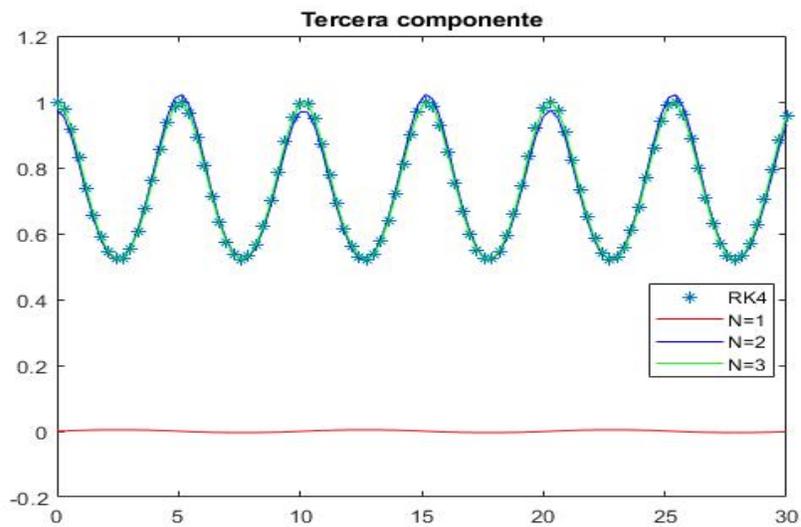


Figura 3.5: Aproximación de la tercera componente y_3 .

Aparentemente podemos observar que la primera y tercera componente se calculan bien considerando solo dos funciones de base. Sin embargo, para el cálculo de la segunda componente es necesario tener, al menos, tres funciones de base. Para ver las distintas aproximaciones en función del número de bases elegidas, en las Figuras 3.6,

3.7 y 3.8 presentamos los errores puntuales cometidos con cada una de las aproximaciones.

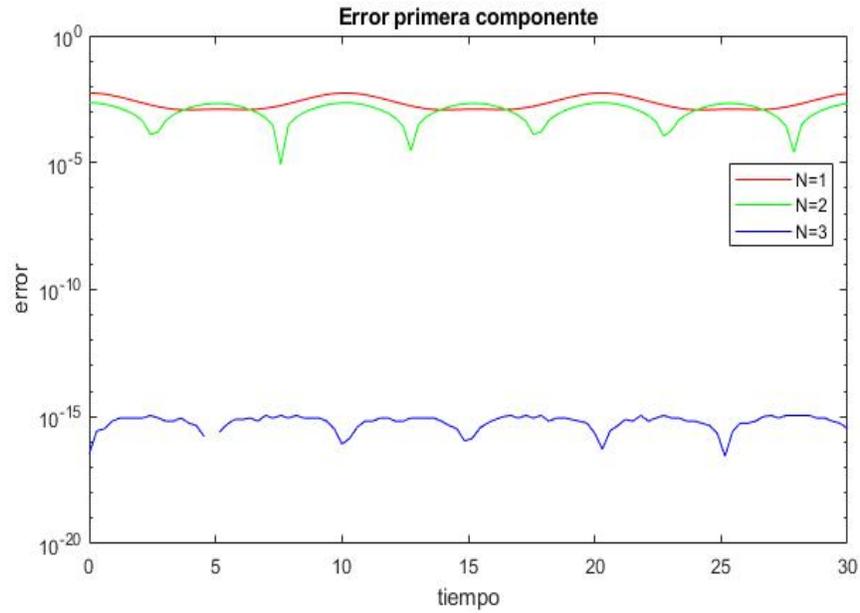


Figura 3.6: Error cometido al aproximar la componente y_1 de la solución.

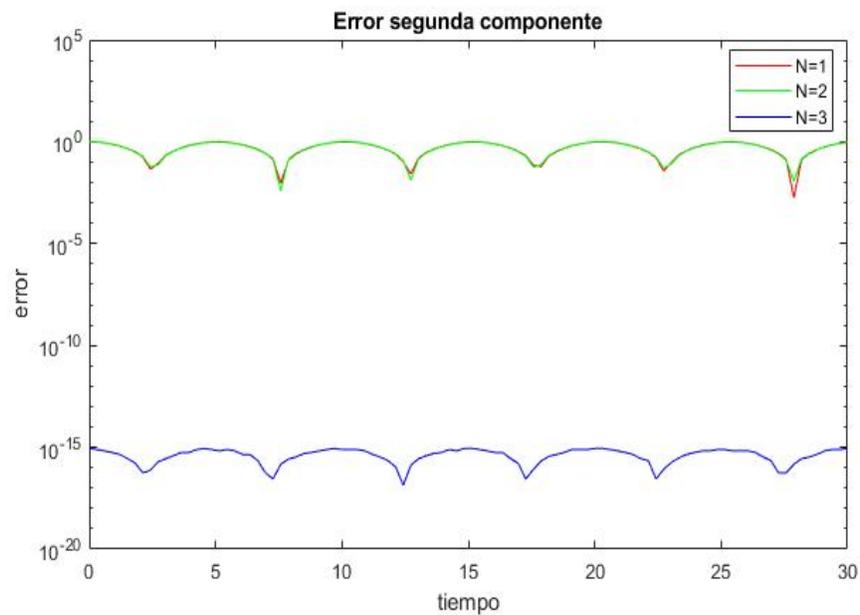


Figura 3.7: Error cometido al aproximar la componente y_2 de la solución.

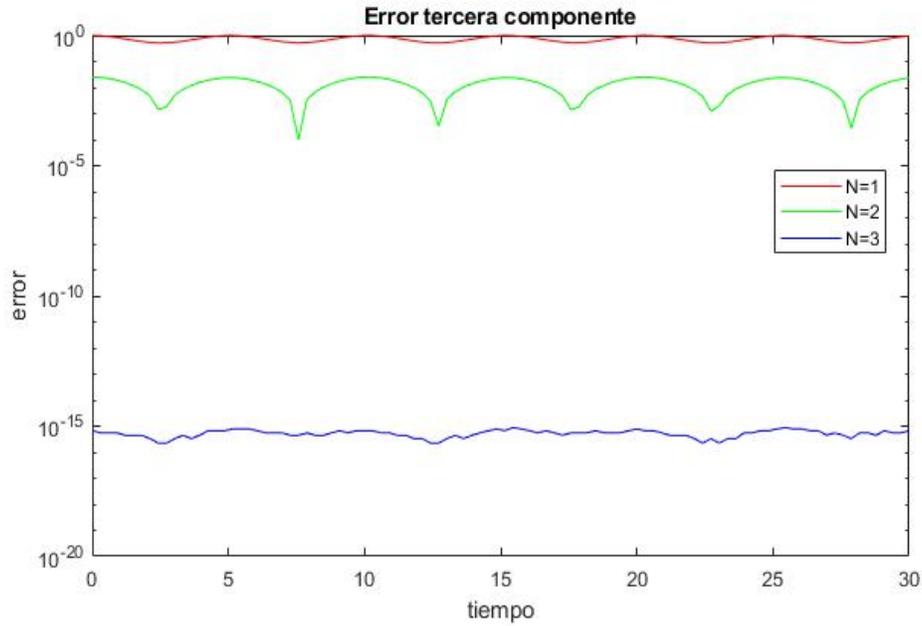


Figura 3.8: Error cometido al aproximar la componente y_3 de la solución.

Como podemos observar en las gráficas anteriores, con una combinación lineal de tres funciones de base obtenemos una buena aproximación con un error cometido de alrededor de 10^{-15} , cerca del ϵ de la máquina.

Para terminar, vamos a añadir unas tablas donde podamos comparar la aproximación que obtenemos para algunos tiempos concretos dependiendo del número de funciones de base escogidas. Si pensamos en que el tiempo lo estamos midiendo en días, vamos a calcular las aproximaciones a los 10, 15, 20 y 25 días, aproximadamente.

10 días	Valor aproximado	N=1	N=2	N=3
y_1	-0.42218	-0.41691	-0.42005	-0.42218
y_2	0.93498	-8.92247e-04	-0.02383	0.93498
y_3	0.95297	-0.00142	0.92934	0.95297

Tabla 2: Aproximación de la solución a los 10 días según el número de funciones de base tomadas.

15 días	Valor aproximado	N=1	N=2	N=3
y_1	0.34912	0.35038	0.34704	0.34912
y_2	-0.95578	7.49848e-04	-0.02364	-0.95578
y_3	0.96785	0.00119	0.99082	0.96785

Tabla 3: Aproximación de la solución a los 15 días según el número de funciones de base tomadas.

20 días	Valor aproximado	N=1	N=2	N=3
y_1	-0.54455	-0.53949	-0.54252	-0.54455
y_2	0.88935	-0.00115	-0.02333	0.88935
y_3	0.92059	-0.00184	0.89810	0.92059

Tabla 4: Aproximación de la solución a los 20 días según el número de funciones de base tomadas.

25 días	Valor aproximado	N=1	N=2	N=3
y_1	0.47732	0.47856	0.47533	0.47732
y_2	-0.91582	0.00102	-0.02262	-0.91582
y_3	0.93929	0.00163	0.96130	0.93929

Tabla 5: Aproximación de la solución a los 25 días según el número de funciones de base tomadas.

Bibliografía

- [1] BARNSTON, A. G. & ROPELEWSKI, C. F., Prediction of ENSO episodes using canonical correlation analysis, *Journal of Climate* 5, 1992, 1316–1345.
- [2] BAYLY, P. V., JOHNSON, E. E., WOLF, P. D., SMITH, W. M., & IDEKER, R. E., Predicting patterns of epicardial potentials during ventricular fibrillation, *IEEE Transactions on Biomedical Engineering* 42, 1995, 898–907.
- [3] CHACÓN, T., FERNÁNDEZ, U. & GÓMEZ M., Computational Modeling of Gurney Flaps and Microtabs by POD Method, *Energies* 11(8), 2018.
- [4] EPUREANU, B. I., HALL, K. C., & DOWELL, E. H., Reduced-order models of unsteady viscous flows in turbomachinery using viscous-inviscid coupling, *Journal of Fluids and Structures* 15, 2001, 255–273.
- [5] GOLUB, G.H. & VAN LOAN, C.F., *Matrix Computations*, The Johns Hopkins University Press, Baltimore, 1990.
- [6] GOHBERG, I. & GOLDBERG, S., *Basic operator theory*, Birkhäuser, 1980.
- [7] LEEN, T. K., RUDNICK, M., & HAMMERSTROM, R., Hebbian feature discovery improves classifier efficiency, in *Proceedings of the IJCNN*, IEEE, Piscataway, NJ, 1990, 51–56.
- [8] LIANG, Y. C., LEE, H. P., LIM, S. P., LIN, W. Z., LEE, K. H., & WU, C. G., Proper orthogonal decomposition and its applications, Part I: Theory', *Journal of Sound and Vibration* 252, 2002, 527–544.
- [9] MEES, A. I., RAPP, P. E., & JENNINGS, L. S., Singular value decomposition and embedding dimension, *Physical Review A* 36, 1987, 340–346.

- [10] NOCEDAL, J. & WRIGHT, S.J, Numerical Optimization, Springer Series in Operation Research, 2006.
- [11] SIROVICH, L., Turbulence and the dynamics of coherent structures, Quarterly of Applied Mathematicss, XLV, 1987, 561-590.
- [12] VOLKWEIN, S., Proper Orthogonal Descomposition: Theory and Reduced-Order Modelling, Lecture Notes, University of Konstanz, 2013.
- [13] WAX, M. & KAILATH, T., Detection of signals by information theoretic criteria, IEEE Transactions on Acoustics, Speech and Signal Processing 33, 1985, 387–392.
- [14] WU, C. G., LIANG, Y. C., LIN, W. Z., LEE, H. P., & LIM, S. P., A note on equivalence of proper orthogonal decomposition methods, Journal of Sound and Vibration 265, 2003, 1103–1110.