mathematical methods - week 2

Eigenvalue problems

Georgia Tech PHYS-6124

Homework HW #2 due Tuesday, September 1, 2020

== show all your work for maximum credit,
== put labels, title, legends on any graphs
== acknowledge study group member, if collective effort
== if you are LaTeXing, here is the exerWeek2.tex

Exercise 2.1 *Three masses on a loop* 8 points
Exercise 2.2 *Examples of singular value decomposition* 2 points + 3 bonus points

Total of 10 points = 100 % score. Bonus points accumulate, can help you later if you miss a few problems.

edited May 15, 2021

Week 2 syllabus

Tuesday, August 25, 2020

> If I had had more time, I would have written less
> — Blaise Pascal, a remark made to a correspondent

Tuesday's lecture is related to AWH Chapter 6 *Eigenvalue problems* (click here). The fastest way to watch any week's lecture videos is by letting YouTube run

- *the course playlist*

- Please do not get intimated by the length of this week's notes - they are here more for me than for you, as notes on these topics for future reference. If you understand the online lectures and can solve the exercises, you are good. The notes you can quickly skim over...
  - Sect. 2.2 *Using symmetries*
  - Sect. 2.3 *Normal modes*: The free vibrations of systems, for undamped systems with total energy conserved for which the frequencies of oscillation are real.
  - *Normal modes*
  - Example 2.1 *Vibrations of a classical CO_2 molecule*
  - *A Hamiltonian with a symmetry (4:46 min)*
  - *CO_2 molecule (4:07 min)*
  - *Projection operators (5:33 min)*
  - *(Anti)symmetric subspaces (3:04 min)*
  - *Zero mode (5:19 min)*
  - AWH Example 6.2.3 *Degenerate eigenproblem*
  - AWH Example 6.5.2 *Normal modes*

- Matrix decompositions in data science
  - Sect. 2.4 *Singular Value Decomposition*
  - *Matrices: physics vs data science*
  - *Singular value decomposition (SVD)*
  - *SVD sample calculation*

2.1 Other sources

Normal modes are important in aeronautical and mechanical engineering (optional reading for week 2, not required for this course):

- MIT 16-07-dynamics is a typical mathematical methods in engineering course. Normal modes are discussed here.

- Example 2.2 pen & paper derivation of normal-modes of the ring of N asymmetric pairs of oscillators (from Gutkin lecture notes example 5.1 C_n *symmetry*).
- Srdjan Ostojić @ostojic_srdjan writes: The singular value decomposition (SVD) course by @eigensteve is great: "These lectures go into depth on the singular value decomposition (SVD), one of the most widely used algorithms for data processing, reduced-order modeling, and high-dimensional statistics, following Chapter 1 of *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control* by Brunton and Kutz [2], with databookuw website and chapters."
- We like the discussion of norms, least square problems, and differences between singular value and eigenvalue decompositions in Trefethen and Bau [4], cited in sect. 2.4.1.
- Andrew: In *Understanding SVD* Reza Bagheri develops SVD step-by-step, starting with the concept of eigenvalues through eigenvalue decomposition and then to SVD. I found it good to review some of the linear algebra I had forgotten. It is long, but it takes time to develop each concept which is a style I find very helpful.
- If you later need SVD in your research, Cline and Dhillon [3] *Computation of the singular value decomposition* seems to be a handy cookbook.
- *Eigen Grandito* - `u/cactus`'s *Principal Components Analysis of the Taco Bell menu*, an NumPy SVD exploration of the *Onion* (1998) classic on Taco Bell's revolutionary Grandito

📖 ChaosBook sect. 6.1 explains the geometrical intuition behind matrix decompositions.

📖 In ChaosBook remark 6.1. *Lyapunov exponents are uncool* Predrag claims that SVD is the wrong thing in dynamics.

- If instead, bedside crocheting is your thing, click here.

2.2 Using symmetries

The big idea #1 of this is week is *symmetry*.

If our physical problem is defined by a (perhaps complicated) Hamiltonian $\mathbf{H}$, another matrix $\mathbf{M}$ (hopefully a very simple matrix) is a symmetry if it commutes with the Hamiltonian

$$[\mathbf{M}, \mathbf{H}] = 0\,. \tag{2.1}$$

Than we can use the spectral decomposition (1.24) of $\mathbf{M}$ to block-diagonalize $\mathbf{H}$ into a sum of lower-dimensional sub-matrices,

$$\mathbf{H} = \sum_i \mathbf{H}_i\,, \qquad \mathbf{H}_i = P_i \mathbf{H} P_i\,, \tag{2.2}$$

and thus significantly simplify the computation of eigenvalues and eigenvectors of **H**, the matrix of physical interest.

2.3 Normal modes

> The big idea #2 of this is week is : *many body systems* (molecules, neuronal networks, ...) are ruled by *collective modes*, not individual particles (atoms, neurons, ...).

In the linear, harmonic oscillator approximation, the classical dynamics of a molecule is governed by the Hamiltonian

$$H = \sum_{i=1}^{N} \frac{m_i}{2} \dot{x}_i^2 + \frac{1}{2} \sum_{i,j=1}^{N} x_i^\top V_{ij} x_j \,,$$

where $\{x_i\}$ are small deviations from the equilibrium, resting points of the molecules labelled i. V_{ij} is a symmetric matrix, so it can be brought to a diagonal form by an orthogonal transformation, to a set of N uncoupled harmonic oscillators or *normal modes* of frequencies $\{\omega_i\}$.

$$x \to y = Ux, \qquad H = \sum_{i=1}^{N} \frac{m_i}{2} \left(\dot{y}_i^2 + \omega_i^2 y_i^2 \right) \,. \tag{2.3}$$

2.4 Singular Value Decomposition

> Everybody knows that the SVD is the best matrix decomposition !!!
>
> — @Daniela_Witten, 21 July 2020

Daniela's Twitter lecture (tweaked by Predrag): If you are in statistics or data science, SVD is the #1 matrix decomposition, and likely the only one you will ever need. And believe me: you are going to need it.

In data science, one often deals with a very large data set X that can be laid out as an rectangular array, vertically arbitrarily high (n time measurements $x_{1j}, x_{2j}, \cdots, x_{nj}$; n faces), and horizontally relatively short (m neuronal voltages $x_{k1}, x_{k2}, \cdots, x_{km}$; m facial features).

What does the SVD do? You give me an $[n \times m]$, $n \geq m$ rectangular matrix X, and I'll give you back 3 matrices, an $[n \times m]$ rectangular matrix U, a diagonal $[m \times m]$ matrix Σ, and unitary $[m \times m]$ matrix V that together "decompose" the matrix X:

$$X = U \Sigma V^T \,. \tag{2.4}$$

U and V are orthogonal matrices (if X is complex, unitary matrices),

$$U^T U = V^T V = V V^T = I_{[m \times m]} \tag{2.5}$$

Σ is diagonal with nonnegative and decreasing elements:

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_m \geq 0\,. \tag{2.6}$$

Some terminology: the diagonal elements of Σ are the *singular values*, and the columns of U and V are the *left* and *right singular vectors* u_k, v_j. Multiply (2.4) from right by V. This implies that if I have the singular value σ_j and the short $[m \times 1]$ singular vector v_j, I can multiply it with my data array X to compute the tall, $[n \times 1]$ singular vector u_j to

$$X\, v_j = \sigma_j u_j\,. \tag{2.7}$$

How do I compute the singular eigenvalues? From (2.4) it follows that V is a rotation that diagonalizes the symmetric $X^T X = \{\sum_j x_{kj} x_{j\ell}\}$ correlation matrix

$$X^T X = V \Sigma^2 V^T\,. \tag{2.8}$$

So this gives us v_k and σ_k (one always picks the positive root of σ_k^2) which we label by the decreasing eigenvalues convention (2.6), and evaluate u_j using (2.7). Why σ_k^2? Rectangular matrix X is dimensionally a strange beast; it relates bricks to oranges, and it's transpose returns oranges to bricks. The result is an (hyper)ellipsoid, with singular vectors as semiaxes, and singular values as lengths along the semiaxes.

Simple as that. What makes this decomposition special (and unique) is the particular set of properties of U, Σ, and V.

Don't be fooled tho: $UU^T \neq I_{n\times n}$!!!!!!! In layperson's terms, the columns of U and V are special: the squared elements of each column of U and V sums to 1, and also the inner product (dot product) between each pair of columns in U equals 0. And the inner product between each pair of columns of V equals 0.

First of all, let's marvel that this decomposition is not only possible, but easily computable, and even unique (up to sign flips of columns of U and V). Like, why on earth should *every matrix* X be decomposable in this way?

Magic, that's why. OK, so, its existence is magic. But, is it also useful? Well, YES.

Suppose you want to approximate X with a pair of vectors: that is, a rank-1 approx. Well, the world's best rank-1 approximation to X, in terms of residual sum of squares, is given by the first columns of U and V:

$$X \approx \sigma_1 u_1 v_1^T \tag{2.9}$$

is literally the best you can do!!

OK, but what if you want to approximate X using two pairs of vectors (a rank-2 approximation)? Just calculate

$$X \approx \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T\,, \tag{2.10}$$

and call it a day.

Want an even better approximation, using rank-k? You *literally can't beat this one*

$$X \approx \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \ldots + \sigma_k u_k v_k^T \tag{2.11}$$

so please don't bother trying.

OK, so, the SVD gives me the best possible way to approximate any matrix. What is this good for??!!

Ever heard of *principal components analysis* (PCA)? This is just the SVD (after centering columns of X to have mean 0). Columns of V are PC loading vectors. Columns of U (up to scaling) are PC score vectors.

Bam!!!

How about if you want to impute missing values in your data matrix X? (In finance 'impute' means "to assign (a value) to something by inference from the value of the products to which it contributes.) Assuming that the elements of X are randomly missing (rather than, for instance, larger elements more likely to be missing), the SVD gives you an effective and easy way to impute those values!!

First, fill in missing elements using (say) the column mean. Then, compute SVD to get rank-k approx for (e.g.) $k = 3$. Replace the missing elements with the elements of rank-k approx. Rinse and repeat until your answer stops changing.

Voila!! I am not making this up!! The SVD is like a matrix X-ray. For instance:

$$X^T X = V\Sigma^2 V^T\,, \quad (X^T X)^{-1} = V\Sigma^{-2}V^T \tag{2.12}$$

$$X(X^T X)^{-1}X^T = UU^T\,. \tag{2.13}$$

Take a minute to breathe this in. Formulas (2.12), (2.13) have no U's, and the 3rd (hat matrix from least squares) has no V's or Σ's!

Wowzers!

Now, you may ask "well what about the eigen-decomposition?" Well I can dispense with that concern in 1 tweet. A symmetric matrix A (the only type worth eigen-decomposing, IMO - internetese for "in my opinion") is just $A = X^T X$ for some X. And $X^T X = V\Sigma^2 V^T$ by (2.12). SVD $\gg$ eigen-decomposition. QED

Not convinced? Need me to spell it out for u? Singular vectors are the eigenvectors of A, and singular values are the square roots of the eigenvalues!!! So, the SVD gives you the eigen-decomposition for free!!! Eigen-decomposition *just got owned by the* SVD. That's how it's done!!!!!

The SVD is a great 1-stop shop for data analysis.

Need to know if X is multi-collinear, before fitting least squares? Check out the singular values. If σ_1/σ_m is huge then least squares is a bad idea.

If $n > m$, or if $\sigma_m = 0$ then bad news bears, X isn't even invertible!! To know if the matrix $X^T X$ is invertible, you just have to check whether the smallest singular value is non-zero. Want to know the rank of X? It's just # of non-zero singular values!

Want the Moore-Penrose pseudo-inverse (though please be careful — there are better ways to approximate a matrix inverse)? That's basically the rank-k approx from earlier, but with $1/\sigma_k$ instead of σ_k!

And please don't troll me with your comments about how you prefer the QR or LU decompositions. I'm a working mom with 3 kids at home in the midst of a pandemic, I know you don't mean it, and I literally don't have time for this. (@SusCrockford, March 12, 2020 concurs: "The next academic dude who posts about how much work Isaac Newton or whoever got done at Cambridge during the plague I'm coming over to

their house with my snotty 4y and staying there to develop my genius while he deals with the kid and then we'll see who discovers laws of nature.")

And if you prefer a specialty matrix decomposition, like NMF, then I've got news for you: you got fooled, because that's just a souped up SVD. Honest to god. If you remain convinced that the NMF or any other decomposition discovered in the past 80 years can hold a candle to the SVD, then I can get you a great price on a Tesla-branded vegan unicorn made out of CRISPR. I'll send it to you as soon as you give me all your Bitcoin.

The SVD is super magical and there's so much I've left unsaid. While you can compute it using a single line of code in R or any other halfway decent programming language, it's fun, easy, and safe to DIY with matrix multiplies!!!!

I hope that this thread has helped to grow your appreciation for this magical decomposition. The more you learn about the SVD, the more you will love it. It will take you far on your statistics and data science journey. Godspeed.

The rest is on YouTube.

2.4.1 Eigen values vs. singular values

> It's now more important to learn boring algebra than to practice fun rock throwing. So you take your choice. If you choose happiness over survival too consistently–well, then you die happy. Or else, you thrive grumpily. It's the tragedy of the human condition. About time we changed it, in my humble opinion.
>
> — Hans Moravec

Trefethen and Bau [4] *Numerical Linear Algebra*:

> $\cdots$ not all matrices (even square ones) have an eigenvalue decomposition [here week 1 and sect. 2.3], but all matrices (even rectangular ones) have a singular value decomposition [here sect. 2.4]. In applications, eigenvalues tend to be relevant to problems involving the behavior of iterated forms of A, such as matrix powers A^k or exponentials $\exp A$, whereas singular vectors tend to be relevant to problems involving the magnitudes of elements of A , or its inverse.

2.4.2 SVD in dynamical systems

Tosif Ahamed @_mlechha : There are lots of applications in recovering dynamics from data. I'd also like to plug our own eigen-worms, using SVD to go from experimental observations of moving worms to their periodic orbits (and more) arXiv:1911.10559.

On 2020-08-20 graduate student Daniel Dylewsky, U Washington, gave a good presentation *Koopman Approximations for Multiscale Nonlinear Physics using Dynamic Mode Decomposition*, based on Daniel Dylewsky, Eurika Kaiser, Steven L. Brunton and J. Nathan Kutz *Principal Component Trajectories (PCT): Nonlinear dynamics as a superposition of time-delayed periodic orbits* arXiv:2005.14321.

2.4.3 SVD in rocket science

Steven L. Brunton *et al. Data-Driven Aerospace Engineering: Reframing the Industry with Machine Learning* arXiv:2008.10740, a review: "$\cdots$ The aerospace industry is poised to capitalize on big data and machine learning, which excels at solving the types of multi-objective, constrained optimization problems that arise in aircraft design and manufacturing. Indeed, emerging methods in machine learning may be thought of as data-driven optimization techniques that are ideal for high-dimensional, non-convex, and constrained, multi-objective optimization problems, and that improve with increasing volumes of data."

2.4.4 SVD in theoretical neuroscience

Srdjan Ostojić again: During my physics education, I have never heard of singular value decomposition.

Almost all matrices in physics are symmetric, and in that case SVD reduces to eigenvalue decomposition.

But for non-symmetric, or especially non-square matrices, SVD is the fundamental tool. So over the recent years, part of the theoretical neuroscience community has been rediscovering how useful SVD is.

For instance, basic results on perceptrons can be understood in a simple way using SVD. Dynamics of learning in deep networks can be understood based on SVD: arXiv:1312.6120, on Pnas, arXiv:1809.10374. Non-linear dynamics in recurrent neural networks can be analyzed by starting from the SVD of the connectivity matrix, and keeping dominant terms: arXiv:1711.09672, bioRxiv:350801v3, arXiv:2007.02062, arXiv:1909.04358, bioRxiv:2020.07.03.185942v1. Non-normal transient dynamics in recurrent networks: on sciencedirect, arXiv:1811.07592.

References

[1] G. B. Arfken, H. J. Weber, and F. E. Harris, *Mathematical Methods for Physicists: A Comprehensive Guide*, Seventh (Academic, New York, 2013).

[2] S. L. Brunton and J. N. Kutz, *Data-driven Science and Engineering: Machine Learning, Dynamical Systems, and Control* (Cambridge Univ. Press, Cambridge UK, 2019).

[3] A. K. Cline and I. S. Dhillon, "Computation of the singular value decomposition", in *Handbook of Linear Algebra* (CRC Press, 2006), pp. 45-1–45-13.

[4] L. N. Trefethen and D. Bau, *Numerical Linear Algebra* (SIAM, 1997).

2.5 Examples

Example 2.1. ***Vibrations of a classical CO_2 molecule:*** *Consider one carbon and two oxygens constrained to the x-axis [1] and joined by springs of stiffness k, as shown*

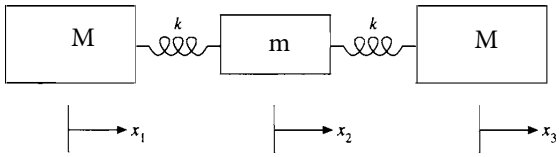

Figure 2.1: A classical colinear CO_2 molecule [1].

in figure 2.1. Newton's second law says

$$\begin{aligned} \ddot{x}_1 &= -\frac{k}{M}(x_1 - x_2) \\ \ddot{x}_2 &= -\frac{k}{m}(x_2 - x_3) - \frac{k}{m}(x_2 - x_1) \\ \ddot{x}_3 &= -\frac{k}{M}(x_3 - x_2)\,. \end{aligned} \tag{2.14}$$

The normal modes, with time dependence $x_j(t) = x_j \exp(it\omega)$, *are the common frequency* ω *vibrations that satisfy (2.14),*

$$\mathbf{H}\mathbf{x} = \begin{pmatrix} A & -A & 0 \\ -a & 2\,a & -a \\ 0 & -A & A \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \omega^2 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\,, \tag{2.15}$$

where $a = k/m$, $A = k/M$. *Secular determinant* $\det(\mathbf{H} - \omega^2 \mathbf{1}) = 0$ *now yields a cubic equation for* ω^2.

You might be tempted to stick this $[3\times 3]$ *matrix into* `Mathematica` *or whatever, but please do that in some other course. What would understood by staring at the output? In this course we think.*

First thing to always *ask yourself is: does the system have a* symmetry*? Yes! Note that the* CO_2 *molecule (2.14) of figure 2.1 is invariant under* $x_1 \leftrightarrow x_3$ *interchange, i.e., coordinate relabeling by matrix* σ *that commutes with our law of motion* $\mathbf{H}$,

$$\sigma = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}\,, \qquad \sigma\mathbf{H} = \mathbf{H}\sigma = \begin{pmatrix} 0 & -A & A \\ -a & 2\,a & -a \\ A & -A & 0 \end{pmatrix}\,. \tag{2.16}$$

We can now use the symmetry operator σ *to simplify the calculation. As* $\sigma^2 = \mathbf{1}$, *its eigenvalues are* ± 1, *and the corresponding symmetrization, anti-symmetrization* projection operators *(1.30) are*

$$P_+ = \frac{1}{2}(\mathbf{1} + \sigma)\,, \qquad P_- = \frac{1}{2}(\mathbf{1} - \sigma)\,. \tag{2.17}$$

The dimensions $d_i = \operatorname{tr} P_i$ *of the two subspaces are*

$$d_+ = 2\,, \qquad d_- = 1\,. \tag{2.18}$$

As σ *and* $\mathbf{H}$ *commute, we can now use spectral decomposition (1.24) to block-diagonalize* $\mathbf{H}$ *to a 1-dimensional and a 2-dimensional matrix.*

On the 1-dimensional antisymmetric subspace, the trace of a $[1\times 1]$ *matrix equals its sole matrix element equals it eigenvalue*

$$\lambda_- = \mathbf{H}P_- = \frac{1}{2}(\operatorname{tr}\mathbf{H} - \operatorname{tr}\mathbf{H}\sigma) = (a + A) - a = \frac{k}{M}\,,$$

so the corresponding eigenfrequency is $\omega_-^2 = k/M$. To understand its physical meaning, write out the antisymmetric subspace projection operator (2.18) explicitly. Its non-vanishing columns are proportional to the sole eigenvector

$$P_- \;=\; \frac{1}{2}\begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix} \;\Rightarrow\; \mathbf{e}^{(-)} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} . \tag{2.19}$$

In this subspace the outer oxygens are moving in opposite directions, with the carbon stationary.

On the 2-dimensional symmetric subspace, the trace yields the sum of the remaining two eigenvalues

$$\lambda_+ + \lambda_0 = \mathrm{tr}\,\mathbf{H}P_+ = \frac{1}{2}(\mathrm{tr}\,\mathbf{H} + \mathrm{tr}\,\mathbf{H}\sigma) = (a + A) + a = \frac{k}{M} + 2\,\frac{k}{m} .$$

We could disentangle the two eigenfrequencies by evaluating $\mathrm{tr}\,\mathbf{H}^2 P_+$, for example, but thinking helps again.

There is still another, translational symmetry, so obvious that we forgot it; if we change the origin of the x-axis, the three coordinates $x_j \to x_j - \delta x$ change, for any continuous translation δx, but the equations of motion (2.14) do not change their form,

$$\mathbf{H}\,\mathbf{x} \;=\; \mathbf{H}\,\mathbf{x} + \mathbf{H}\,\delta\mathbf{x} \;=\; \omega^2\mathbf{x} \quad\Rightarrow\quad \mathbf{H}\,\delta\mathbf{x} \;=\; 0\,. \tag{2.20}$$

So any translation $\mathbf{e}^{(0)} = \delta\mathbf{x} = (\delta x, \delta x, \delta x)$ is a nul, 'zero mode' eigenvector of $\mathbf{H}$ in (2.16), with eigenvalue $\lambda_0 = \omega_0^2 = 0$, and thus the remaining eigenfrequency is $\omega_+^2 = k/M + 2\,k/m$. As we can add any nul eigenvector $\mathbf{e}^{(0)}$ to the corresponding $\mathbf{e}^{(+)}$ eigenvector, there is some freedom in choosing $\mathbf{e}^{(+)}$. One visualization of the corresponding eigenvector is the carbon moving opposite to the two oxygens, with total momentum set to zero.

(Taken from AWH Example 6.2.3 Degenerate eigenproblem*, but done here using symmetries.)*

Example 2.2. ***Vibrational spectra of molecules:*** *Consider the ring of pair-wise interactions of two kinds of molecules sketched in figure 2.2 (a), given by the potential*

$$V(z) = \frac{1}{2}\sum_{i=1}^{N}\left(k_1(x_i - y_i)^2 + k_2(x_{i+1} - y_i)^2\right) , \qquad z_i = \begin{pmatrix} x_i \\ y_i \end{pmatrix} , \tag{2.21}$$

whose $[2N\times 2N]$ matrix form is (aside to the cognoscenti: this is a Toeplitz matrix):

$$V_{ij} = \frac{1}{2}\begin{pmatrix} k_1+k_2 & -k_1 & 0 & 0 & 0 & \dots & 0 & 0 & -k_2 \\ -k_1 & k_1+k_2 & -k_2 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & -k_2 & k_1+k_2 & -k_1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & -k_1 & k_1+k_2 & -k_2 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & -k_2 & k_1+k_2 & -k_1 \\ -k_2 & 0 & 0 & 0 & 0 & \dots & 0 & -k_1 & k_1+k_2 \end{pmatrix}$$

This potential matrix is a holy mess. How do we find an orthogonal transformation (2.3) that diagonalizes it? Look at figure 2.2 (a). Molecules lie on a circle, so that suggests we should use a Fourier representation. As the $i = 1$ labelling of the starting molecule

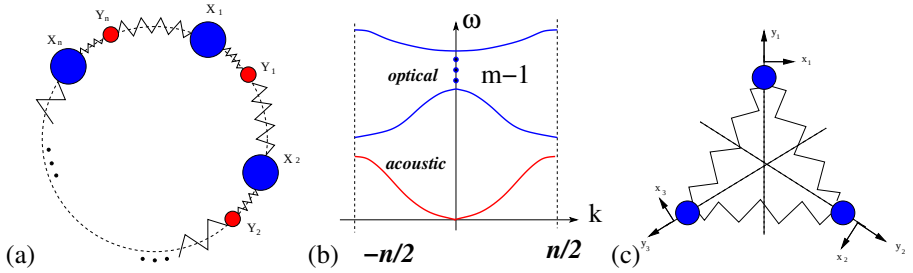

Figure 2.2: (a) Chain with circular symmetry. (b) Dependance of frequency on the representation wavenumber k. (c) Molecule with D_3 symmetry. (B. Gutkin)

on a ring is arbitrary, we are free to relabel them, for example use the next molecule pair as the starting one. This relabelling is accomplished by the $[2N\times 2N]$ *permutation matrix (or 'one-step shift', 'stepping' or 'translation' matrix)* M *of form*

$$\underbrace{\begin{pmatrix} 0 & 0 & \dots & 0 & I \\ I & 0 & \dots & 0 & 0 \\ 0 & I & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & I & 0 \end{pmatrix}}_{M} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} z_n \\ z_1 \\ z_2 \\ \vdots \\ z_{n-1} \end{pmatrix}, \qquad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad z_i = \begin{pmatrix} x_i \\ y_i \end{pmatrix} \tag{2.22}$$

Projection operators corresponding to M *are worked out in example 10.1. They are* N *distinct* $[2N\times 2N]$ *matrices,*

$$P_k = \begin{pmatrix} I & \bar\lambda I & \bar\lambda^2 I & \dots & \bar\lambda^{N-2} I & \bar\lambda^{N-1} I \\ \lambda I & I & \bar\lambda I & \dots & \bar\lambda^{N-3} I & \bar\lambda^{N-2} I \\ \lambda^2 I & \lambda I & I & \dots & \bar\lambda^{N-4} I & \bar\lambda^{N-3} I \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \lambda^{N-2} I & \lambda^{N-3} I & \lambda^{N-4} I & \dots & I & \bar\lambda I \\ \lambda^{N-1} I & \lambda^{N-2} I & \lambda^{N-2} I & \dots & \lambda I & I \end{pmatrix}, \qquad \lambda = \exp\left(\frac{2\pi i}{N} k\right) \tag{2.23}$$

which decompose the $2N$*-dimensional configuration space of the molecule ring into a direct sum of* N *2-dimensional spaces, one for each discrete Fourier mode* $k = 0, 1, 2, \cdots, N-1$.

The system (2.21) is clearly invariant under the cyclic permutation relabelling M*,* $[V, M] = 0$ *(though checking this by explicit matrix multiplications might be a bit tedious), so the* P_k *decompose the interaction potential* V *as well, and reduce its action to the* k*th 2-dimensional subspace. Thus the* $[2N\times 2N]$ *diagonalization (2.3) is now reduced to a* $[2\times 2]$ *diagonalization which one can do by hand. The resulting* k*th space is spanned*

by two $2N$-dimensional vectors, which we guess to be of form:

$$\eta_1 = \frac{1}{\sqrt{n}} \begin{pmatrix} 1 \\ 0 \\ \lambda \\ 0 \\ \vdots \\ \lambda^{n-1} \\ 0 \end{pmatrix}, \qquad \eta_2 = \frac{1}{\sqrt{n}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ \lambda \\ \vdots \\ 0 \\ \lambda^{n-1} \end{pmatrix}.$$

In order to find eigenfrequences we have to consider action of V on these two vectors:

$$V\eta_1 = (k_1 + k_2)\eta_1 - (k_1 + k_2\lambda)\eta_2\,, \qquad V\eta_2 = (k_1 + k_2)\eta_2 - (k_1 + k_2\bar{\lambda})\eta_1\,.$$

The corresponding eigenfrequencies are determined by the equation:

$$\begin{aligned} 0 &= \det\left(\begin{pmatrix} k_1 + k_2 & -(k_1 + k_2\lambda) \\ -(k_1 + k_2\bar{\lambda}) & k_1 + k_2 \end{pmatrix} - \frac{\omega^2}{2} I\right) \qquad \Longrightarrow \\ \frac{1}{2}\omega_{\pm}^2(k) &= k_1 + k_2 \pm |k_1 + k_2\lambda^k|\,, \end{aligned} \tag{2.24}$$

one acoustic ($\omega(0) = 0$), one optical, see figure 2.2 (b) and the acoustic and optical phonons wiki. *(B. Gutkin)*

Example 2.3. ***An SVD hand calculation.*** *Given a rectangular $[n \times m] = [4 \times 3]$ "data matrix"*

$$X = \begin{pmatrix} 0 & \sqrt{2} & 0 \\ 1 & 2 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \tag{2.25}$$

in this example we implement by hand its singular value decomposition

$$X = U\Sigma V^T\,. \tag{2.26}$$

A side remark: by inspection, the 1st and 3rd rows of X are not independent from the 2nd, the rank of the data matrix X is 2, so expect one zero eigenvalue.

The 'right', $[m \times m] = [3 \times 3]$ correlation matrix (see (2.7) and (2.8)) is

$$C_r = X^T X = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 6 & 2 \\ 2 & 2 & 2 \end{pmatrix}. \tag{2.27}$$

The zeroes of its characteristic polynomial

$$\det(C_r - \lambda 1) = (-8 + \lambda)(-2 + \lambda)\lambda = 0 \tag{2.28}$$

yield eigenvalues

$$\{\lambda_1, \lambda_2, \lambda_3\} = \{8, 2, 0\} \tag{2.29}$$

You are free to find the corresponding eigenvectors any way you like. If you use projection operators, you will also need the matrix squared:

$$C_r^2 = \begin{pmatrix} 12 & 20 & 12 \\ 20 & 44 & 20 \\ 12 & 20 & 12 \end{pmatrix}$$

The associated projection operators are:

$$P_1 = \frac{(C_r - 2\cdot 1)(C_r - 0\cdot 1)}{(8-2)(8-0)} = \frac{1}{6\cdot 8}(C_r^2 - 2C_r) = \frac{1}{6}\begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix} \tag{2.30}$$

$$P_2 = \frac{(C_r - 8\cdot 1)(C_r - 0\cdot 1)}{(2-8)(2-0)} = \frac{1}{6\cdot 2}(-C_r^2 + 8C_r) = \frac{1}{3}\begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & - & 1 \end{pmatrix}$$

$$P_3 = \frac{(C_r - 8\cdot 1)(C_r - 2\cdot 1)}{(0-8)(0-2)} = \frac{1}{8\cdot 2}(C_r^2 - 10C_r + (8\cdot 2)\,1) = \frac{1}{2}\begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

Each column of a projection operator is the same right eigenvector, with a different prefactor, and its rows are likewise proportional to the same left eigenvector. SVD, however, demands that the eigenvectors be normalized to unit length, for example

$$v_1 = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \tag{2.31}$$

The three normalized right singular vectors, taken as the columns, form the rotation matrix

$$V = \left(\begin{array}{c|c|c} v_1 & v_2 & v_3 \end{array}\right) = \left(\begin{array}{c|c|c} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} \\ \frac{2}{\sqrt{6}} & -\frac{2}{\sqrt{3}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \end{array}\right) \tag{2.32}$$

The "left", "big" correlation matrix

$$C_l = X\,X^T = \begin{pmatrix} 2 & 2\sqrt{2} & 0 & 0 \\ 2\sqrt{2} & 6 & 2 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{2.33}$$

characteristic polynomial has the same non-zero eigenvalues

$$\det(C_l - \lambda 1) = (-8+\lambda)(-2+\lambda)\lambda^2 = 0\,, \tag{2.34}$$

but an extra zero eigenvalue. Going through the same algebra as for C_r, we find that C_l (unnormalized) eigenvectors can be presented as columns of matrix

$$\hat{U} = \begin{pmatrix} \sqrt{2} & 3 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \sqrt{2} & -1 & 1 & 0 \end{pmatrix} \tag{2.35}$$

After normalization to unit length we refer to them as the left singular vectors.

The singular values are, by definition, the positive square roots of C_r or C_l eigenvalues

$$\{\sigma_1, \sigma_2, \sigma_3\} = \{2\sqrt{2}, \sqrt{2}, 0\} \tag{2.36}$$

so the diagonal singular values matrix is given by

$$\Sigma = \begin{pmatrix} 2\sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{2.37}$$

SVG decomposition (2.26) is of form $X = U\Sigma V^T$. Now that we have the right eigenvectors matrix V, and the diagonal singular values matrix Σ, we could compute the left eigenvectors matrix $U = XV/\Sigma$, as can be done in some of the examples in exercise 2.2. But zero singular values make this a bit tricky, so here we compute instead also U from the C_l eigenvalue equation, and verify that we indeed get the rectangular data matrix exercise 2.25 back:

$$U\Sigma V^T = \begin{pmatrix} 0 & \sqrt{2} & 0 \\ 1 & 2 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

A sanity check: Mathematica does all this in one line:
$\{U, \Sigma, V\} =$ ***SingularValueDecomposition***$[X] \quad \Rightarrow$

$$U = \begin{pmatrix} \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{\sqrt{3}}{2} & 0 & 0 & -\frac{1}{2} \\ \frac{1}{2\sqrt{3}} & \sqrt{\frac{2}{3}} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 \end{pmatrix}, \qquad \Sigma = \begin{pmatrix} 2\sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{2.38}$$

$$V = \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} \\ \sqrt{\frac{2}{3}} & -\frac{1}{\sqrt{3}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \end{pmatrix} \tag{2.39}$$

verifying (2.32), (2.37), and (2.35).

Exercises

2.1. **Three masses on a loop.** Three identical masses, connected by three identical springs, are constrained to move on a circle hoop as shown in figure 2.3. Find the normal modes. Hint: write down coupled harmonic oscillator equations, guess the form of oscillatory solutions. Then use basic matrix methods, i.e., find zeros of a characteristic determinant, find the eigenvectors, etc.. (Kimberly Y. Short)

2.2. **Examples of singular value decomposition.** Bring, by hand calculation, the following matrices into SVD form:

$$\begin{aligned} A &= \begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix}, \; B = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}, \; C = \begin{pmatrix} 0 & 2 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \\ D &= \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \; E = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}. \end{aligned} \tag{2.40}$$

The goal is to verify that any matrix, including these, has the unique SVD decomposition, and (2.8), (2.6) and (2.7) should suffice for the job.

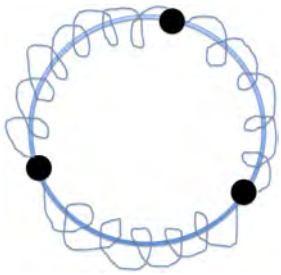

Figure 2.3: Three identical masses are constrained to move on a hoop, connected by three identical springs such that the system wraps completely around the hoop. Find the normal modes.