Homework 4

AM213A

Kevin Silberberg

2025-02-24

Part 1

Let us plot the data for the ‘atkinson.dat’ file.

```
using GLMakie, DelimitedFiles

# read data 'atkinson.dat' from file
data = readdlm("../atkinson.dat")
x = data[:, 1]
y = data[:, 2]

# plot data points for part 1
fig = Figure()
ax = Axis(fig[1, 1], title = "points read from atkinson.dat")
lines!(ax, x, y, color = :blue, linestyle = (:dash, :dense))
save("part1.png", fig)
```

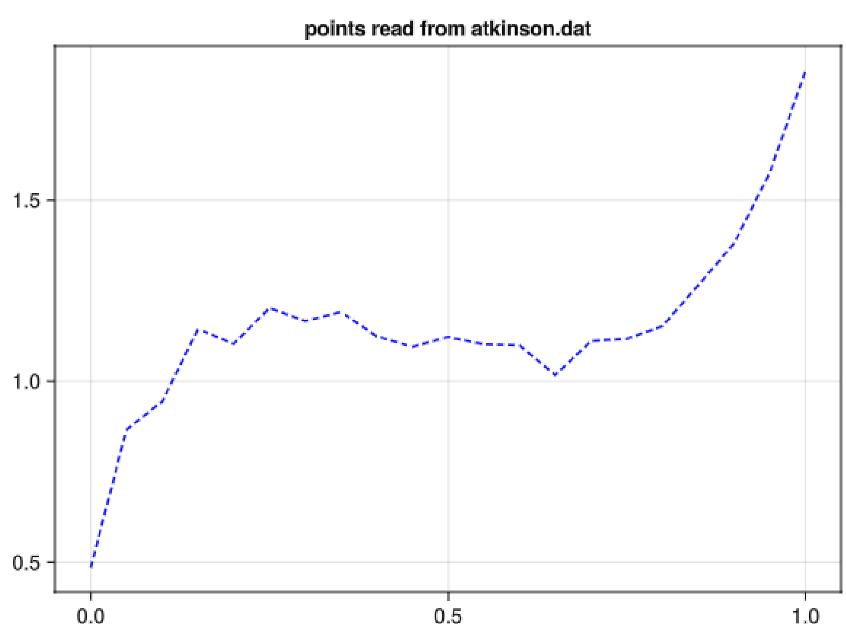

Figure 1: Y by X plot of points read from file ‘atkinson.dat’

Cholesky solution of the least squares problem

For the points plotted above,

$$\mathbf{P} = \begin{bmatrix} x_0 & y_0 \\ x_1 & y_1 \\ \vdots & \\ x_m & y_m \end{bmatrix} \tag{1}$$

we want to solve the following least squares problem,

$$\mathbf{Vb} = \mathbf{y} \tag{2}$$

where $\mathbf{V}$ is a Vandermonde matrix such that the $\mathbf{x}$ column of $\mathbf{P}$ in the vandermonde matrix is

$$\mathbf{V} = \begin{bmatrix} | & | & | & | & | \\ \mathbf{x}^0 & \mathbf{x}^1 & \mathbf{x}^2 & \cdots & \mathbf{x}^n \\ | & | & | & | & | \end{bmatrix} \tag{3}$$

from (2) we can solve the least squares problem by left multiplying by $\mathbf{V}^T$ on both sides

$$\mathbf{V}^T\mathbf{Vb} = \mathbf{V}^T\mathbf{y} \tag{4}$$

where $\mathbf{V}^T\mathbf{V}$ is the Gram-Matrix of $\mathbf{V}$. Since $\mathbf{V}^T\mathbf{V}$ is a square symmetric positive definite matrix we can perform a cholesky decomposition on the $\mathbf{V}^T\mathbf{V}$ to obtain

$$\mathbf{V}^T\mathbf{V} = \mathbf{LL}^T \tag{5}$$

so that we have

$$\mathbf{LL}^T\mathbf{b} = \mathbf{V}^T\mathbf{y} \tag{6}$$

then we apply forward substitution to find an unknown $\mathbf{z}$ where we use the lower triangular matrix $\mathbf{L}$,

$$\mathbf{Lz} = \mathbf{V}^T\mathbf{y} \tag{7}$$

finally we solve for $\mathbf{b}$, the coefficients, using backward substitution

$$\mathbf{L}^T\mathbf{b} = \mathbf{z} \tag{8}$$

once we have the solution vector $\mathbf{b}$ we can obtain an approximation for $\mathbf{y}$ by fitting a curve of polynomial order n given by the column vector $\mathbf{Vb}$.

we can then find how well this curve fit the original points by calculating the frobenius norm of the error given by

$$\|\mathbf{y} - \mathbf{Vb}\|_F \tag{9}$$

cholesky.c, cholesky.h

cholesky function

The function cholesky takes in a matrix, checks if A is square then proceeds with the algorithm on page 52/53 of the notes. The algorithm loops over the columns, calculates a new diagonal element and checks if the diagonal element is not less than zero before taking the square root. Then it calculates the elements below the diagonal and checks if the entries are symmetric for all elements before proceeding to the next index.

choleskyBacksub function

The function choleskyBacksub takes in the matrix that was overwritten by the cholesky decomposition and performs forward substitution to find the unknown column vector $\mathbf{z}$ in (7).

Then backward substitution is performed to find the solution vector $\mathbf{b}$ in (8) by iterating backwards from the last row to the first and checking if the diagonals are less than a small number (here we choose to use 1×10^{-12}) before dividing by that diagonal element. In our function we chose to have intermediate data structures and then finally overwrite the solution vector into the input vector and deallocate all intermediate matrices before returning.

mainCholesky

In the driver program for the first problem, the function takes in two arguments, the filename of the points to read and an unsigned integer number representing the polynomial order to create the vandermonde matrix.

The driver program

- Reads the points into an $n \times 2$ matrix $\mathbf{P}$.
- Copies the points into column vectors $\mathbf{x}$ and $\mathbf{y}$.
- Creates the vandermonde matrix $\mathbf{V}$ with n degree polynomials
- Constructs the Gram-Matrix $\mathbf{V}^T\mathbf{V}$
- Left matrix multiplies $\mathbf{V}^T\mathbf{y}$
- Compute the cholesky decomposition of $\mathbf{V}^T\mathbf{V}$
- Performs forward and backsubstitution to find $\mathbf{x} = (\mathbf{V}^T\mathbf{V})^{-1}\mathbf{V}^T\mathbf{y}$
- Prints the solution vector $\mathbf{x}$
- Calculates and prints the error by taking the Frobenius norm of the vector (9)
- Write the fitted points $\mathbf{V}\mathbf{x} = \mathbf{y}$ to file and prints the file name.

Fit data with 3rd degree polynomial using single precision floating point arithmetics

```
make clean
make single
./main cholesky atkinson.dat chopolysinglethree.dat 3
```

The polynomial coefficients are

$$y = f(x) = 7.668485x^3 - 11.127914x^2 + 4.725732x + 0.57467 \quad (10)$$

The frobenius norm error on the fit is:

$$\|\mathbf{y} - \mathbf{V}\mathbf{x}\|_F = 0.192745 \tag{11}$$

The following is a comparison of the fitted curve vs the data points

```
# read the polynomial line of fit from file
# cholesky ploynomial order 3
Vb = readdlm("../chopolysinglethree.dat")
fig = Figure()
ax = Axis(fig[1, 1], title ="cholesky solution, polynomial order 3")
scatter!(ax, x, y, color = :blue, label = "original points")
lines!(ax, x, Vb[:, 1], color = :red, label = "3rd order polyfit")
Legend(fig[1, 2], ax)
save("part1cholesky3.png", fig)
```

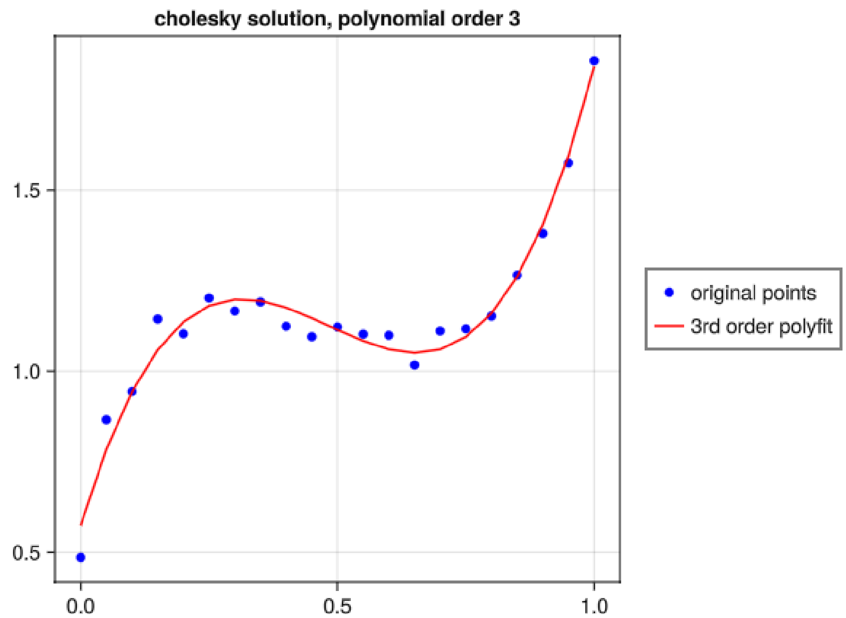

Figure 2: Vb by X plot of scattered points read from file 'atkinson.dat' and line of fit created with single precision and polynomial order 3

Fit data with 5th degree polynomial using single precision floating point arithmetics

```
./main cholesky atkinson.dat chopolysinglefive.dat 5
```

The polynomial coefficients are

$$y = f(x) = 17.164490x^5 - 45.097218x^4 + 49.775738x^3 - 27.587082x^2 + 7.089328x + 0.512521 \tag{12}$$

The frobenius norm error on the fit is:

$$\|\mathbf{y} - \mathbf{V}\mathbf{x}\|_F = 0.140591 \tag{13}$$

The following is a comparison of the fitted curve vs the data points

```
# read the polynomial line of fit from file
# cholesky ploynomial order 5
Vb = readdlm("../chopolysinglefive.dat")
fig = Figure()
ax = Axis(fig[1, 1], title ="cholesky solution, polynomial order 5")
scatter!(ax, x, y, color = :blue, label = "original points")
lines!(ax, x, Vb[:, 1], color = :red, label = "5th order polyfit")
Legend(fig[1, 2], ax)
save("part1cholesky5.png", fig)
```

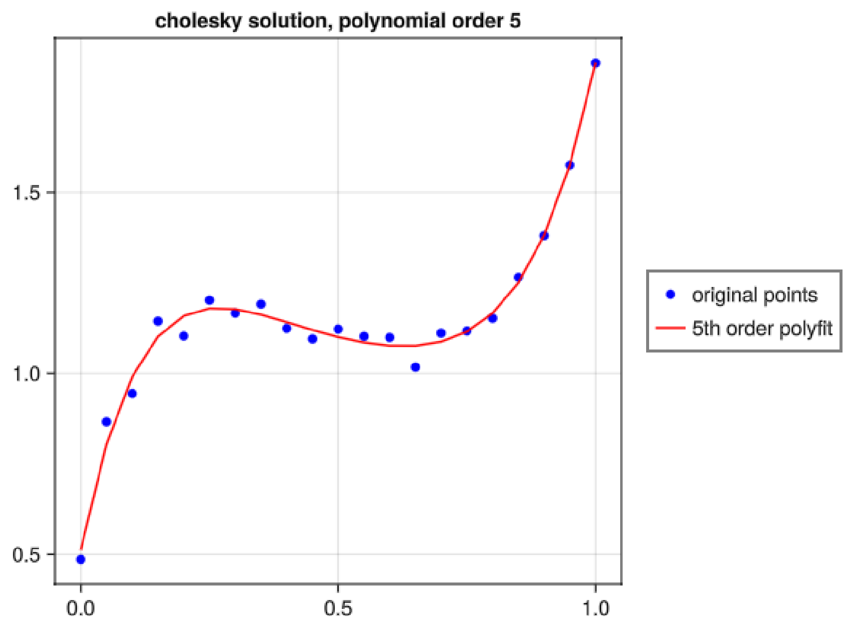

Figure 3: Vb by X plot of scattered points read from file 'atkinson.dat' and line of fit created with single precision and polynomial order 5

For the given data, the maximum degree of the polynomial that should fit the data is order three. This is because we want the most parsimonious line of fit for the data without over fitting. If we plotted the frobenius norm error for increasing polynomial order, we should expect that the 'elbow plot' produced should identify $n = 3$ as the polynomial order that reduces the error the most.

```
# elbow plot
using LinearAlgebra
vandermonde(x::Vector{Float64}, n::Int)::Matrix{Float64} = hcat([x.^i for i in 0:n]...)

error = Vector{Float64}(undef, 6)
for n in eachindex(error)
    V = vandermonde(x, n)
    VTV = V' * V
    VTy = V' * y
    b = VTV \ VTy
    ŷ  = V * b
    error[n] = norm(y - ŷ )
end
```

```
fig = Figure()
ax = Axis(fig[1, 1],
          title = "Frobenius norm error by increasing polynomial order",
          xlabel = "n",
          ylabel = "error",
          xticks = eachindex(error))
lines!(ax, error, color = :red)
save("part1elbowplot.png", fig)
```

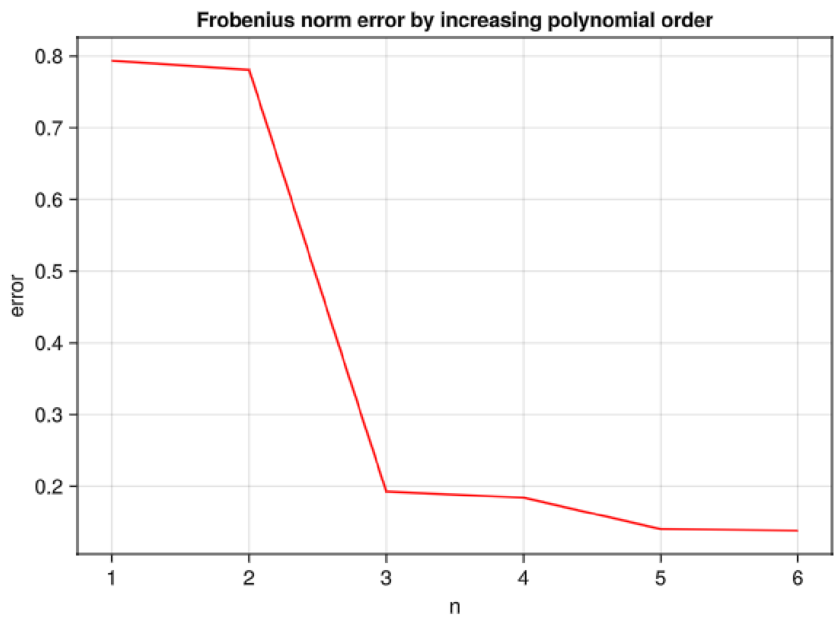

Figure 4: Elbow plot of the Frobenius norm error for the line of best fit for increase polynomial order. We can clearly see that $n = 3$ is the point at which the error drops the most and increasing polynomial orders only marginally improve the error.

The algorithm begins to fail when n or the polynomial order of the Vandermonde matrix gets too large. This is mainly becuase the original vandermonde matrix becomes ill conditioned due to the larger polynomial columns having either vary large or vary small terms. Additionall the Gram-Matrix becomes less and less positive definite as you increase the polynomial order. For my code this happens at anything greater than $n > 5$ because I have a condition in the cholesky decomposition algorithm that checks if the matrix is positive definite and the threshold is set to 1×10^{-12}.

QR solution of the least-squares problem

House householer decomposition

Given a problem

$$\mathbf{A}\mathbf{x} = \mathbf{b} \tag{14}$$

The householder decomposition transforms the matrix $\mathbf{A}$ into $\mathbf{R}$ an upper triangular matrix. It saves the vectors $\mathbf{v}_j$ in the lower trangular portion of $\mathbf{A}$ and saves the diagonal elements of $\mathbf{R}$ into a seperate array. When we want to use the decomposition for solving the system (14), we construct $\mathbf{Q}$ by iteratively computing

$$\mathbf{Q} = \left(\mathbf{I} - 2\mathbf{v}_j\mathbf{v}_j^T\right)\mathbf{Q} \tag{15}$$

starting from the identity matrix. The householder vectors $\mathbf{v}_j$ are stored in the columns of the modified $\mathbf{A}$ matrix, and each transformation $\mathbf{I} - 2\mathbf{v}_j\mathbf{v}_j^T$ updates $\mathbf{Q}$ for all columns $\mathbf{v}_j$.

Fit atkinson.dat with 3rd degree polynomial, single precision

The following is the command line usage that was used to produce the data that reads the atkinson.dat file, performs QR decomposition on the gram-matrix of the vandermonde matrix and solves the system

$$\mathbf{A}\mathbf{x} = \mathbf{b} \tag{16}$$
$$\mathbf{Q}\mathbf{R}\mathbf{x} = \mathbf{b} \tag{17}$$
$$\mathbf{R}\mathbf{x} = \mathbf{Q}^T\mathbf{b} \tag{18}$$
$$\mathbf{x} = \mathbf{R}^{-1}\mathbf{Q}^T\mathbf{b} \tag{19}$$

where $\mathbf{A} = \mathbf{V}^T\mathbf{V}$ and $\mathbf{b} = \mathbf{V}^T\mathbf{y}$ and $\mathbf{V}$ is the vandermonde matrix described by (3).

the program then computes the polynomial line of fit by matrix vector multiplying the orginal matrix $\mathbf{V}$ by the found coefficients $\mathbf{x}$ and saves it to file for plotting in julia.

```
code$ make clean
code$ make single
code$ ./main householder atkinson.dat housepolysinglethree.dat 3
```

The polynomial coefficients are

$$f(x) = \mathbf{V}\mathbf{x} = 7.668801x^3 - 11.128375x^2 + 4.725902x + 0.574661 \tag{20}$$

The Frobenius norm error is

$$\|\mathbf{y} - \mathbf{V}\mathbf{x}\|_F = 0.192745 \tag{21}$$

Now let us plot the polynomial line of fit that was written to file.

```julia
# read 3rd order polynomial line of fit from file
# householder holder decomp with Gaussian elimination
Vb = readdlm("../housepolysinglethree.dat")
fig = Figure()
ax = Axis(fig[1, 1], title = "householder decomp with gauss elimination, polynomial order 3")
scatter!(ax, x, y, color = :blue, label = "original points")
lines!(ax, x, Vb[:, 1], color = :red, label = "3rd order polyfit")
Legend(fig[1, 2], ax)
save("part1house3.png", fig)
```

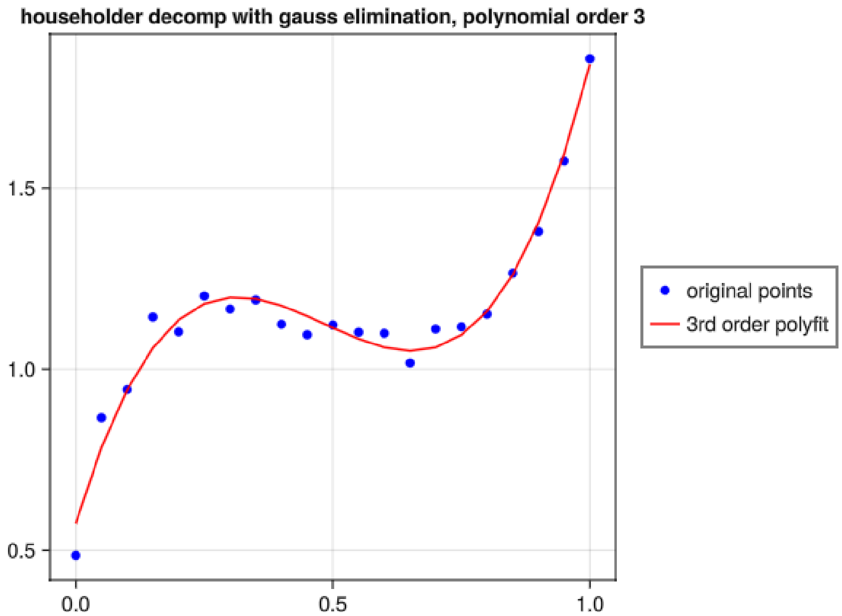

Figure 5: Vb by X plot of scattered points read from file 'atkinson.dat' and line of fit created with single precision and polynomial order 3 via the householder QR decomposition and Gauss elimination for inverting R

Fit atkinson.dat with 5th degree polynomial, single precision

The following produces the same as above except with polynomial order 5

```
./main householder atkinson.dat housepolysinglefive.dat 5
```

The polynomial coefficients are,

$$f(x) = \mathbf{V}\mathbf{x} = 16.666607x^5 - 43.796494x^4 + 48.565464x^3 - 27.113325x^2 + 7.020390x + 0.514441 \quad (22)$$

The Frobenius norm error on the fit is,

$$\|\mathbf{y} - \mathbf{V}\mathbf{x}\|_F = 0.140794 \quad (23)$$

Now let us plot the polynomial line of fit for order 5 that was written to file.

```
# read 5th order polynomial line of fit from file
# householder holder decomp with Gaussian elimination
Vb = readdlm("../housepolysinglefive.dat")
fig = Figure()
ax = Axis(fig[1, 1], title = "householder decomp with gauss elimination, polynomial order 5")
scatter!(ax, x, y, color = :blue, label = "original points")
lines!(ax, x, Vb[:, 1], color = :red, label = "5th order polyfit")
Legend(fig[1, 2], ax)
save("part1house5.png", fig)
```

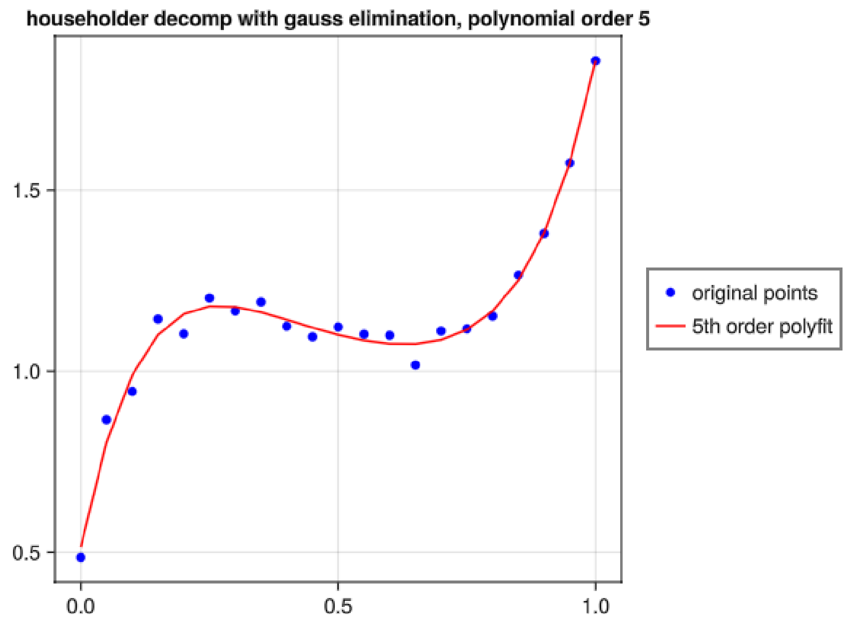

Figure 6: Vb by X plot of scattered points read from file 'atkinson.dat' and line of fit created with single precision and polynomial order 5 via the householder QR decomposition and Gauss elimination for inverting R

At which point does the algorithm fail?

There is no point at which the algorithm fails. We can perform a line of best fit for polynomial of order 100 and the algorithm calculates the coefficients just the same. This is because the householder decomposition is backwards stable and it does not rely the matrix $\mathbf{A}$ being symmetric positive definiite.

Discuss your findings about the Frobenius norms of $\mathbf{A} - \mathbf{QR}$ and $\mathbf{Q}^T\mathbf{Q} - \mathbf{I}$

The frobenius norm of the error for $\mathbf{A} - \mathbf{QR}$ for polynomial order 3 was found to be

$$\|\mathbf{A} - \mathbf{QR}\|_F = 0.000015 \tag{24}$$

and for polynomial order 5 was,

$$\|\mathbf{A} - \mathbf{QR}\|_F = 0.000025 \tag{25}$$

The increase in the error is likely due to the accumulated round off errors from decomposing $\mathbf{A} \to \mathbf{QR}$ in single precision and then performing the matrix multiplication $\mathbf{QR}$. The growth in error from polynomial order $n = 3 \to 5$ is likely proportional to the number of floating point operations between the decompositions for a matrix of size 4×4 vs 6×6.

Part 2

Question 1.

If $\mathbf{P}$ is an orthogonal projector, then $\mathbf{I} - 2\mathbf{P}$ is unitary.

Proof. Let $\mathbf{P}$ be an orthogonal projector.

Recall that an orthogonal projector has the following two properties:

$$\mathbf{P} = \mathbf{P}^2 \tag{26}$$
$$\mathbf{P}^T = \mathbf{P} \tag{27}$$

Let $\mathbf{Q} = \mathbf{I} - 2\mathbf{P}$.

Let us prove that $\mathbf{Q}$ is a unitary matrix by showing that $\mathbf{Q}\mathbf{Q}^T = \mathbf{Q}^T\mathbf{Q} = \mathbf{I}$.

First let us show that $\mathbf{Q} = \mathbf{Q}^T$

$$\mathbf{Q}^T = (\mathbf{I} - 2\mathbf{P})^T = \mathbf{I} - 2\mathbf{P}^T \tag{28}$$

Since $\mathbf{P} = \mathbf{P}^T$ we have that $\mathbf{Q} = \mathbf{Q}^T$. Well since $\mathbf{Q} = \mathbf{Q}^T$, it immediately follows that $\mathbf{Q}\mathbf{Q}^T = \mathbf{Q}^T\mathbf{Q}$. Let us show that $\mathbf{Q}\mathbf{Q}^T = \mathbf{I}$.

$$\mathbf{Q}\mathbf{Q}^T = (\mathbf{I} - 2\mathbf{P})(\mathbf{I} - 2\mathbf{P}) \tag{29}$$
$$= \mathbf{I}^2 - 2\mathbf{I}\mathbf{P} - 2\mathbf{I}\mathbf{P} + 4\mathbf{P}^2 \tag{30}$$
$$= \mathbf{I} - 4\mathbf{P} + 4\mathbf{P}^{\mathbf{2}} \tag{31}$$
$$= \mathbf{I} \tag{32}$$

□

Geometric Interpretation

The orthogonal projector $\mathbf{P}$ maps any vector $\mathbf{v}$ onto a subspace $\mathbf{S}$ such that the vector $\mathbf{v}$ can be decomposed as,

$$\mathbf{v} = \mathbf{P}\mathbf{v} + (\mathbf{I} - \mathbf{P})\mathbf{v} \tag{33}$$

where $\mathbf{P}\mathbf{v}$ lies in the subspace $\mathbf{S}$ and $(\mathbf{I} - 2\mathbf{P})\mathbf{v}$ lies in the subspace orthogonal to $\mathbf{S}$.

We can analyze the transformation that the unitary matrix $\mathbf{Q} = \mathbf{I} - 2\mathbf{P}$ has on $\mathbf{v}$ by comparing the decomposition of $\mathbf{v}$ and $\mathbf{Q}\mathbf{v}$.

$$\mathbf{Q}\mathbf{v} = (\mathbf{I} - 2\mathbf{P})\mathbf{v} \tag{34}$$
$$= (\mathbf{I} - \mathbf{P} - \mathbf{P})\mathbf{v} \tag{35}$$
$$= \mathbf{v} - \mathbf{P}\mathbf{v} - \mathbf{P}\mathbf{v} \tag{36}$$
$$= -\mathbf{P}\mathbf{v} + (\mathbf{v} - \mathbf{P}\mathbf{v}) \tag{37}$$
$$= -\mathbf{P}\mathbf{v} + (\mathbf{I} - \mathbf{P})\mathbf{v} \tag{38}$$

Comparing this to (33), we can see that $\mathbf{Q}$ applies a reflection about the subspace onto which $\mathbf{P}$ projects.

Question 2.

Consider the matrix

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}. \tag{39}$$

What is the orthogonal projector $\mathbf{P}$ onto range($\mathbf{A}$), and what is the image under $\mathbf{P}$ of $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

Solution

$\mathbf{P}$ is given by,

$$\begin{aligned} \mathbf{P} &= \mathbf{A}\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A^T} && (40) \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} && (41) \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix} && (42) \\ &= \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} && (43) \end{aligned}$$

The image under $\mathbf{P}$ of $\mathbf{v}$ is thus,

$$\begin{aligned} \mathbf{Pv} &= \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} && (44) \\ &= \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} && (45) \end{aligned}$$

Question 3.

Let $\mathbf{P} \in \mathbb{R}^{m\times m}$ be a nonzero projector.

Part a

The 2-norm of the matrix $\mathbf{P}$ is greater than or equal to one.

$$\|\mathbf{P}\|_2 \geq 1 \tag{46}$$

Proof. Let $\mathbf{P} \in \mathbb{R}^{m\times m}$ and $\mathbf{P} \neq \mathbf{0}$ be a projector that satisfies the property that $\mathbf{P} = \mathbf{P}^2$.

By the property of matrix norms we have,

$$\|\mathbf{PP}\|_2 \leq \|\mathbf{P}\|_2 \|\mathbf{P}\|_2 \tag{47}$$
$$\|\mathbf{P}\|_2 \leq \|\mathbf{P}\|_2 \|\mathbf{P}\|_2 \tag{48}$$
$$\|\mathbf{P}\|_2 \geq \frac{\|\mathbf{P}\|_2}{\|\mathbf{P}\|_2} = 1 \tag{49}$$

□

If $\mathbf{P}$ is an orthogonal projector then $\|\mathbf{P}\|_2 = 1$

Proof. Given that $\mathbf{P}$ is an orthogonal projector, for any vector $\mathbf{x} \neq 0$,

$$\|\mathbf{x}\|^2 = \|\mathbf{Px} + (\mathbf{I} - \mathbf{P})\,\mathbf{x}\|^2 \quad \text{where } \mathbf{Px} \perp (\mathbf{I} - \mathbf{P})\,\mathbf{x} \tag{50}$$
$$\|\mathbf{x}\|^2 = \|\mathbf{Px}\|^2 + \|\,(\mathbf{I} - \mathbf{P})\,\mathbf{x}\|^2 \quad \text{by the pythagorean theorem} \tag{51}$$

if (51) is true, then the following must also be true.

$$\|\mathbf{x}\|^2 \geq \|\mathbf{Px}\|^2 \tag{52}$$
$$\|\mathbf{x}\| \geq \|\mathbf{Px}\| \tag{53}$$
$$\frac{\|\mathbf{x}\|}{\|\mathbf{x}\|} \geq \frac{\|\mathbf{Px}\|}{\|\mathbf{x}\|} \tag{54}$$
$$1 \geq \frac{\|\mathbf{Px}\|}{\|\mathbf{x}\|} \tag{55}$$

if (55) is true for all $\mathbf{x}, \mathbf{P}$ then the following must also be true.

$$1 \geq \sup_{\mathbf{x} \in \mathbb{R}^m} \frac{\|\mathbf{Px}\|}{\|\mathbf{x}\|} \tag{56}$$

finally, by the definition of the generalized p-norm, we have

$$1 \geq \|\mathbf{P}\| \tag{57}$$

since in the first part we proved that $\|\mathbf{P}\| \geq 1$, therefore we have equality when $\mathbf{P}$ is an orthogonal projector. □

Part b

If $\mathbf{P}$ is an orthogonal projector, then $\mathbf{P}$ is positive semi-definite with its eigenvalues are either zero or 1.

Proof. Let $\mathbf{P}$ be an orthogonal projector, satisfying the property that $\mathbf{P}^T = \mathbf{P}$, and $\mathbf{P}^2 = \mathbf{P}$.

Let $\mathbf{x} \in \mathbb{R}^m$. Recall that the definition for a matrix $\mathbf{A}$ to be considered positive semi-definite (PSD) is

$$\mathbf{x}^T \mathbf{A} \mathbf{x} \geq 0 \quad \forall \mathbf{x} \tag{58}$$

starting from the left hand side of the definition (58).

$$\begin{aligned}
\mathbf{x}^T\mathbf{P}\mathbf{x} &= \mathbf{x}^T\mathbf{P}^2\mathbf{x} && (59)\\
&= \mathbf{x}^T\mathbf{P}\mathbf{P}\mathbf{x} && (60)\\
&= \mathbf{x}^T\mathbf{P}^T\mathbf{P}\mathbf{x} && (61)\\
&= (\mathbf{P}\mathbf{x})^T\mathbf{P}\mathbf{x} && (62)\\
&= \|\mathbf{P}\mathbf{x}\|_2^2 \geq 0 && (63)
\end{aligned}$$

Now let us show that the eigenvalues of any non-zero projector $\mathbf{P}$ can only take on the values either zero or one.

Let $\lambda, v \neq 0$ be an eigenvalue eigenvector pair of the matrix $\mathbf{P}$

$$\begin{aligned}
\mathbf{P}\mathbf{v} &= \lambda\mathbf{v} && (64)\\
(\mathbf{P}\mathbf{v})^2 &= (\lambda\mathbf{v})^2 && (65)\\
\mathbf{P}^2\mathbf{v} &= \lambda^2\mathbf{v} && (66)\\
\mathbf{P}\mathbf{v} &= \lambda^2\mathbf{v} && (67)\\
\lambda\mathbf{v} &= \lambda^2\mathbf{v} && (68)\\
\lambda(1-\lambda)\mathbf{v} &= 0 && (69)
\end{aligned}$$

thus, $\lambda_i \in \{0, 1\} \quad \forall \lambda_i \in \sigma(\mathbf{P})$ □

Part c

Assume $\mathbf{P}$ is an orthogonal projector and its eigenvalues are all distinct with algebraic multiplicity of 1. What should be the dimension $m \times m$ of $\mathbf{P}$ (i.e., decide what m is)? (Hint: Use the result from (**b**).

Solution

Since the matrix $\mathbf{P}$ is square, it must have m eigenvalues. Since each eigenvalue is distinct with algebraic multiplicity one, and the sum of all the algebraic multiplicities of each eigenvalue must be m, and part (b) has proved that orthogonal projectors can only take on eigenvalues 0 and 1, the dimension of $\mathbf{P}$ must be exactly $m = 2$, such that $\mathbf{P} \in \mathbb{R}^{2\times 2}$, where the first eigenvalue is 0 and the second eigenvalue is 1.

The eigenvector corresponding to the eigenvalue 1 spans the range of $\mathbf{P}$, and the eigenvector corresponding to the eigenvalue 0 lies in the null space of $\mathbf{P}$.

Part d

Let $\mathbf{P} \in \mathbb{R}^{2\times 2}$ be an orthogonal projector whose entries are all nonzero. Identify all possible choices of $\mathbf{P}$. For each of your constructed $\mathbf{P}$, show that it is positive semi-definite and its eigenvalues are either 0 and 1.

Solution

Let $\mathbf{P} = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$

Let us find the entries of $\mathbf{P}$ such that all the entries are non-zero, and they satisfy the property that $\mathbf{P}^2 = \mathbf{P}$ and $\mathbf{P}^T = \mathbf{P}$.

$$
\begin{aligned}
\mathbf{P}^2 &= \begin{bmatrix} a & b \\ b & d \end{bmatrix} \begin{bmatrix} a & b \\ b & d \end{bmatrix} && (70) \\
&= \begin{bmatrix} a^2 + b^2 & ab + bd \\ ab + db & b^2 + d^2 \end{bmatrix} && (71) \\
&= \begin{bmatrix} a & b \\ b & d \end{bmatrix} = \mathbf{P} && (72)
\end{aligned}
$$

equating both sides we have the following system of equations:

$$
\begin{aligned}
a^2 + b^2 &= a && (73) \\
ab + db &= b && (74) \\
b^2 + d^2 &= d && (75)
\end{aligned}
$$

solving for d in terms of a, we have

$$
\begin{aligned}
ab + db &= b && (76) \\
b\,(a + d) &= b && (77) \\
a + d &= 1 && (78) \\
d &= 1 - a && (79)
\end{aligned}
$$

solving for b in terms of a we have,

$$
\begin{aligned}
a^2 + b^2 &= a && (80) \\
b^2 &= a - a^2 && (81) \\
b &= \pm\sqrt{a - a^2} && (82)
\end{aligned}
$$

Thus the matrix $\mathbf{P}$ can be written only in terms of a,

$$
\mathbf{P} = \begin{bmatrix} a & \pm\sqrt{a - a^2} \\ \pm\sqrt{a - a^2} & 1 - a \end{bmatrix} \tag{83}
$$

Notice that if $a < 0$ or $a > 1$ then we can have complex values in the diagonals, but here we are restricting $\mathbf{P}$ to only have real entries. Thus $0 \leq a \leq 1$.

if we let $a = \sin^2\theta$

$$
\mathbf{P} = \begin{bmatrix} \sin^2\theta & |\sin\theta\cos\theta| \\ |\sin\theta\cos\theta| & \cos^2\theta \end{bmatrix} \tag{84}
$$

for all $\theta \in [0, 2\pi]$

We have specifically chosen $\mathbf{P}$ to satisfy the properties of an orthogonal projector, mainly that $\mathbf{P} = \mathbf{P}^2$ and $\mathbf{P} = \mathbf{P}^T$. Thus by the proof in part b, the matrix has eigenvalues strictly 0 and 1, has rank 1 and trace 1, and is a positive semi-definite matrix.

Question 4.

Determine the (i) eigenvalues, (ii) determinant, and (iii) singular values of a Householder reflector. For the eigenvalues, give a geometric argument as well as an algebraic proof.

Solution

Given the definition of a Householder reflector,

$$\mathbf{H} = \mathbf{I} - 2\mathbf{v}\mathbf{v}^T \tag{85}$$

where $\|\mathbf{v}\| = 1$, $\mathbf{H}$ is both symmetric and an orthogonal matrix such that,

$$\begin{aligned}
\mathbf{H} &= \mathbf{H}^T && (86)\\
\mathbf{H}^T &= \mathbf{H}^{-1} && (87)\\
\mathbf{H}\mathbf{H}^T &= \mathbf{I} && (88)
\end{aligned}$$

We can prove the above properties by simply computing $\mathbf{H}\mathbf{H}^T$

$$\begin{aligned}
\mathbf{H}\mathbf{H}^{\mathbf{T}} &= \left(\mathbf{I} - 2\mathbf{v}\mathbf{v}^T\right)\left(\mathbf{I} - 2\mathbf{v}\mathbf{v}^T\right)^T && (89)\\
&= \left(\mathbf{I} - 2\mathbf{v}\mathbf{v}^T\right)\left(\mathbf{I} - 2\left(\mathbf{v}\mathbf{v}^T\right)^T\right) && (90)\\
&= \left(\mathbf{I} - 2\mathbf{v}\mathbf{v}^T\right)\left(\mathbf{I} - 2\mathbf{v}\mathbf{v}^T\right) && (91)\\
&= \mathbf{I} - 2\mathbf{v}\mathbf{v}^T - 2\mathbf{v}\mathbf{v}^T + 4\left(\mathbf{v}\mathbf{v}^T\right)^2 && (92)
\end{aligned}$$

Since $\left(\mathbf{v}\mathbf{v}^T\right)^2 = \mathbf{v}\mathbf{v}^T\mathbf{v}\mathbf{v}^T$ and $\mathbf{v}^T\mathbf{v} = \|\mathbf{v}\| = 1$ we are left with

$$\mathbf{H}\mathbf{H}^T = \mathbf{I} - 4\mathbf{v}\mathbf{v}^T + 4\mathbf{v}\mathbf{v}^T = \mathbf{I} \tag{93}$$

Given the properties above, consider an eigenvalue eigenvector pair, $\lambda, \mathbf{v}$ where $\mathbf{v} \neq 0$.

$$\begin{aligned}
\mathbf{H}\mathbf{v} &= \lambda\mathbf{v} && (94)\\
\mathbf{H}^T\mathbf{H}\mathbf{v} &= \mathbf{H}^T\lambda\mathbf{v} && (95)\\
\mathbf{v} &= \lambda\mathbf{H}^T\mathbf{v} && (96)\\
&= \lambda\mathbf{H}\mathbf{v} && (97)\\
&= \lambda\lambda\mathbf{v} && (98)\\
\lambda^2 &= 1 && (99)\\
\lambda &= \pm 1 && (100)
\end{aligned}$$

It makes sense that the householder reflector has eigenvalue -1 and 1 because, geometrically $\mathbf{H}$ inverts the component of any vector parallel to $\mathbf{v}$ while leaving the orthogonal components unchanged.

Consider the vector $\mathbf{v}$ that is perpendicular to the hyperplane. Applying $\mathbf{H}$ to $\mathbf{v}$ we have,

$$
\begin{aligned}
\mathbf{Hv} &= (\mathbf{I} - 2\mathbf{v}\mathbf{v}^T)\mathbf{v} && (101)\\
&= \mathbf{v} - 2\mathbf{v}\mathbf{v}^T\mathbf{v} && (102)\\
&= \mathbf{v} - 2\mathbf{v} && (103)\\
&= -\mathbf{v} && (104)
\end{aligned}
$$

thus, $\mathbf{v}$ is an eigenvector with eigenvalue -1.

Consider the vector $\mathbf{u}$ that is orthogonal to the vector $\mathbf{v}$, such that $\mathbf{v}^T\mathbf{u} = 0$.

$$
\begin{aligned}
\mathbf{Hu} &= \left(\mathbf{I} - 2\mathbf{v}\mathbf{v}^T\right)\mathbf{u} && (105)\\
&= \mathbf{u} - 2\mathbf{v}\mathbf{v}^T\mathbf{u} && (106)\\
&= \mathbf{u} && (107)
\end{aligned}
$$

Thus the vector $\mathbf{u}$ is an eigenvector with eigenvalue 1 and any vector that is parallel to the hyperplane that is orthogonal to the unit vector $\mathbf{v}$ is invariant under the transformation $\mathbf{H}$.

Now let us find the determinant of $\mathbf{H}$, starting with (88)

$$
\begin{aligned}
\mathbf{H}\mathbf{H}^T &= \mathbf{I} && (108)\\
\det\left(\mathbf{H}\mathbf{H}^T\right) &= \det\left(\mathbf{I}\right) && (109)\\
\det(\mathbf{H})\det(\mathbf{H}^T) &= 1 && (110)\\
\det(\mathbf{H})^2 &= 1 && (111)\\
\det(\mathbf{H}) &= \pm 1 && (112)
\end{aligned}
$$

The singular values of the householder reflector are the square roots of the eigenvalues of the gram-matrix $\mathbf{H}^T\mathbf{H}$ which is equal to identity.

Well the eigenvalues of the identity matrix is equal to 1 with algebraic multiplicity m. Thus the singular values of the householder reflector are all equal to 1.

This is representative of the fact that a unit circle under the transformation $\mathbf{H}$ is invariant and maps to itself.