Numerical Linear Algebra

Kevin Corcoran

February 27, 2022

Contents

Comp total

44/50

3: Assignment 3

Mon 14 Feb 2022 23:36

1 Part 1: Coding Problems

22/25

1.1 Cholesky solution of the least-squares problem

1.1.1 Explaining Cholesky decomposition

We want to solve the least squares problem for $V^TVx = V^Ty$ for the coefficients of of the polynomial for given data points (x_i, y_i) where V is the Vandermonde matrix. We decompose the normal matrix V^TV into its Cholesky decomposition and transform $\hat{y} = V^Ty$

How is L actually formed? -2

$$
\begin{aligned}
V^TVx &= V^Ty \\
LL^Tx &= \hat{y}
\end{aligned}
$$

Letting $L^Tx = \hat{x}$, we first solve for $\hat{x}$ using forward substitution

$$L\hat{x} = \hat{y}.$$

Then we solve for the solution x using the solution we just found for $\hat{x}$ using back substitution

$$L^Tx = \hat{x}.$$

This decomposition works as long as V^TV is positive definite.

1.1.2 Fitting the data to a third degree polynomial

Using Cholesky decomposition, the resulting polynomial is (approximately)

$$\boxed{p(x) \approx 1.8318 - 5.1696 + 11.2024x^2 - 7.2839x^3}.$$

and the two norm error

$$\boxed{\|Vx - y\|_2 \approx 0.244575337}.$$

Plotting the fitted polynomial to the curve

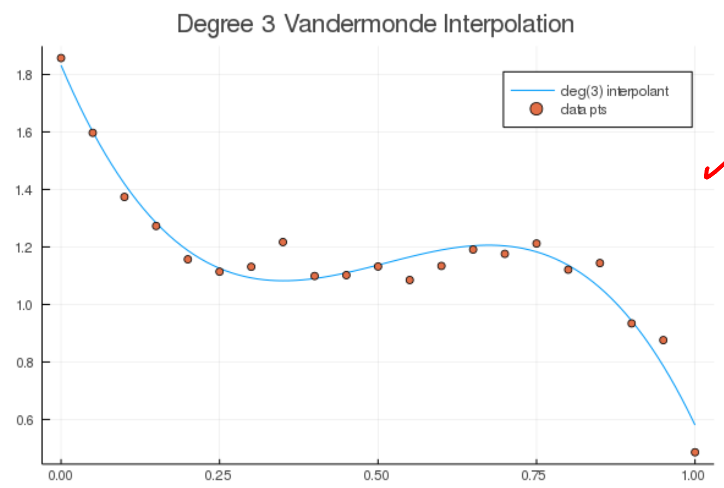

Figure 1: Third degree polynomial

1.1.3 Fitting a fifth degree polynomial

The polynomial

$$\boxed{p(x) \approx 1.8697 - 7.2834 + 28.9995^2 - 59.3135x^3 + 61.7153^4 - 25.4864x^5}.$$

The two norm error

$$\boxed{\|Vx - y\|_2 \approx 0.172760308}.$$

Plotting the polynommial

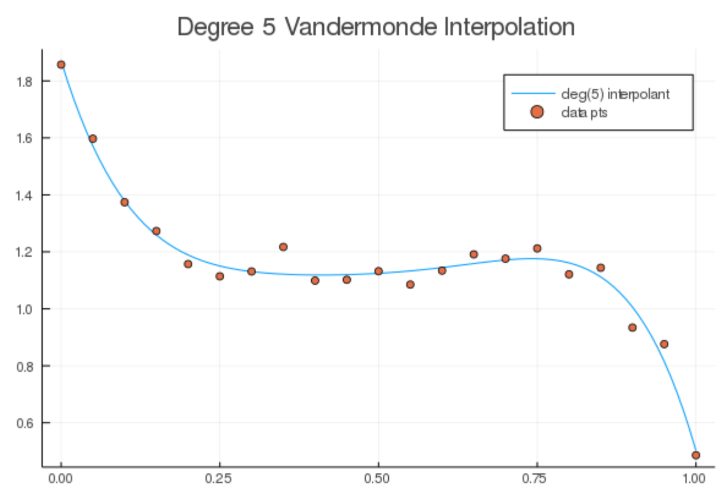

Figure 2: Fifth degree polynomial

They are still fully representable, should comment on importance of relative accuracy.

1.1.4 Maximum degree of the polynomial

what is theoretic max degree?

In single precision, the maximum degree of the polynomial is 7. This is because of round off errors. All of the data points are less than 1, so when we take the 7^{th} power of these numbers inside the Vandermonde matrix, and them multiply them together in V^TV, we end up with values smaller than machine accuracy. Even though the matrix V^TV is not theoretically singular, we end up with a singular matrix with is no longer positive definite.

1.1.5 When does this algorithm fail

This algorithm fails if, for a general matrix A, A^TA is not positive definite, either mathematically or due to floating point errors. We would like to avoid calculating the normal matrix at all.

1.2 QR solution of the least-squares problem

22/25

1.2.1 Explain how Householder QR decomposition works

Householder QR decomposition takes a matrix $A_{m\times n}$ and column by column transforms A into an upper triangular matrix $R_{m\times n}$ using Householder reflectors $H_i = I - 2v_iv_i^T$

How does $H_i A$ behave, and what does it do to column j?

-2

$$H_ia_i = -s_ie_i.$$

where $s_i = \text{sign}(a_{ii})\|a_i\|_2$, and

$$v_i = \frac{a_i + s_ie_i}{\|a_i + s_ie_i\|}.$$

Where the first $i-1$ elements of v_i are 0, and since, from (2.1), H_i is unitary (i.e. $H_i^*H_i = H_i^*H_i = I$) - or orthogonal for H_i real. Then, since the product of orthogonal matrices is orthogonal, we recover the QR decomposition of A

$$H_n \dots H_1A = Q^TA = R.$$

1.2.2 Fitting a third degree polynomial

The polynomial

$$\boxed{p(x) \approx 1.8319 - 5.1705x + 11.2044x^2 - 7.2852x^3}.$$

Two norm error

$$\boxed{\|Vx - y\|_2 \approx 0.244575202}.$$

Plotting the fitted polynomial

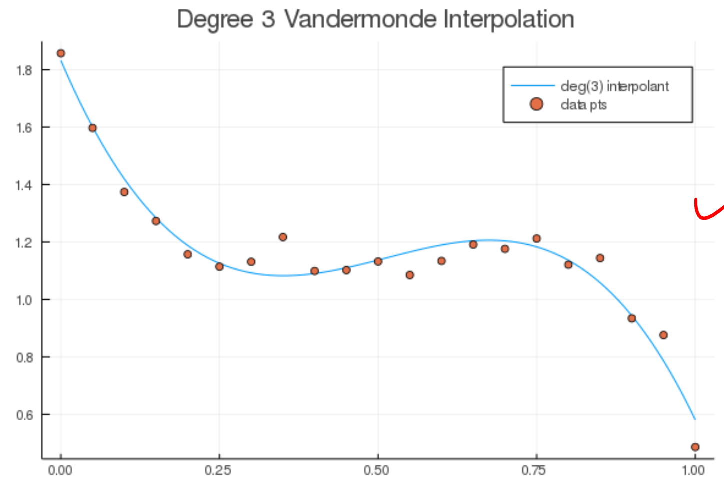

Figure 3: Householder Degree 3

1.2.3 Fitting a fifth degree polynomial

The polynomial

$$\boxed{p(x) \approx 1.8695 - 7.2643x + 28.8177x^2 - 58.7616x^3 + 61.0529x^4 - 25.2123x^5}.$$

Two norm error

$$\boxed{\|Vx - y\|_2 \approx 0.172749385}.$$

Plotting the fitted polynomial

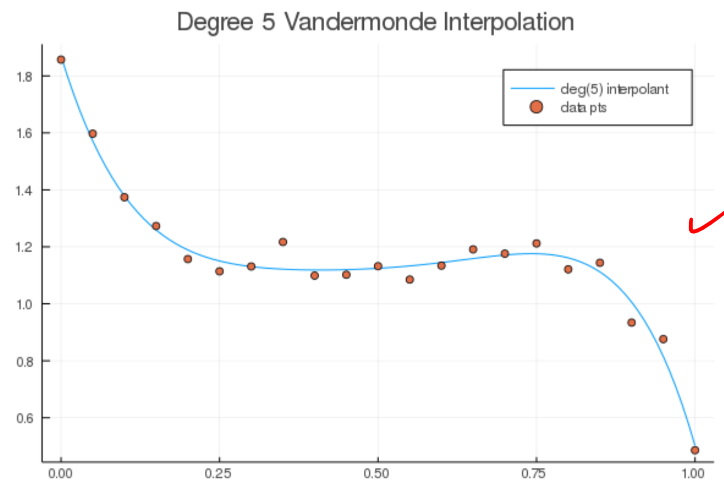

Figure 4: Householder Degree 5

Guarantee is a strong word, but QR is very stable.

1.2.4 When does this algorithm fail

This algorithm guarantees a solution within machine accuracy, so we should be able to find a polynomial of degree equal to the number of data points. (minus 1)

1.2.5 Two norm error of $A - QR$ and $Q^T Q - I$

The error of each column is within single precision machine accuracy. In fact, it is very close to machine accuracy. report them, what is the significance of them? −1

2 Part 2: Theory Problems

2.1 If P is an orthogonal projector, then I − 2P is unitary

We want to show $(I - 2P)^*(I - 2P) = I$. Since P is a projector, $P^2 = P$, and since P is an orthogonal projector, $P^* = P$.

$$\begin{aligned}(I - 2P)^*(I - 2P) &= (I - 2P^*)(I - 2P)\\ &= I - 2P^* - 2P + 4P^*P\\ &= I - 4P + 4P\\ &= I\end{aligned}$$

Therefore, $I - 2P$ is unitary. Geometrically, this is a reflection across the orthogonal compliment $I - P$ of the projection P, and we know reflections are unitary operations since they preserve the norm.

2.2 For $\mathbf{P} \in \mathbb{R}^{\mathbf{m}\times\mathbf{m}}$ a nonzero projector

2.2.1 Show $\|\mathbf{P}\|_2 \geq 1$ with equality if and only if $\mathbf{P}^* = \mathbf{P}$ (an orthogonal projector)

Since P is a projector $P^2 = P$

$$\|P\|_2 = \|P^2\|_2.$$

Since, the 2-norm is submultiplicative (i.e. $\|AB\|_2 \leq \|A\|_2\|B\|_2$), then

$$\begin{aligned}\|P\|_2 &= \|P^2\|_2\\ &= \|PP\|_2\\ &\leq \|P\|_2\|P\|_2\end{aligned}$$

Dividing by $\|P\|_2$

$$\begin{aligned}\frac{\|P\|_2}{\|P\|_2} &\leq \frac{\|P\|_2\|P\|_2}{\|P\|_2}\\ &\Rightarrow \boxed{1 \leq \|P\|_2}\end{aligned}$$

If P is an orthogonal projector, $P^* = P$. We want to show $\|P\|_2 \leq 1$

$$\begin{aligned} \|Px\|_2^2 &= \langle Px, Px \rangle \\ &= \langle x, Px \rangle \qquad P^*P = P^2 = P \\ &\leq \|x\|_2 \|Px\|_2 \qquad \text{By Cauchy-Schwartz} \end{aligned}$$

Dividing by $\|x\|_2\|Px\|_2$, we get

$$\frac{\|Px\|_2}{\|x\|_2} \leq 1.$$

then taking the supremum

$$\sup \frac{\|Px\|_2}{\|x\|_2} = \|P\|_2 \leq 1.$$

This shows that $\boxed{\|P\|_2 = 1}$ when P is an orthogonal projector, since we've shown both $\|P\|_2 \leq 1$, and $\|P\|_2 \geq 1$.

2.2.2 If P is an orthogonal projector, then P is positive semi-definite with its eigenvalues either zero or 1

Since $P = P^*$ is Hermitian, then by the spectral theorem P can be diagonalized by a unitary matrix U ($U^{-1} = U^*$). Hence there exists and eigenbasis for P. Let $(\lambda, \mathbf{v})$ be an eigenpair for P, then

$$\begin{aligned} P\mathbf{v} &= \lambda \mathbf{v} \qquad (1) \\ \Leftrightarrow P^2\mathbf{v} &= \lambda^2 \mathbf{v} \\ P\mathbf{v} &= \lambda^2 \mathbf{v} \qquad P^2 = P \qquad (2) \end{aligned}$$

From equations (1) and (2)

$$\begin{aligned} \lambda &= \lambda^2 \\ \Rightarrow & \boxed{\lambda = 0 \quad \text{or} \quad \lambda = 1} \qquad (3) \end{aligned}$$

From the spectral theorem

$$P = UDU^*$$

Let $\mathbf{y} = U^*\mathbf{x}$ be a change of basis where P is diagonal, then for any $\mathbf{x} \neq 0$

$$\begin{aligned} \mathbf{x}^*P\mathbf{x} &= \mathbf{x}^*UDU^*\mathbf{x} \\ &= \mathbf{y}D\mathbf{y} \\ &= \sum_{i=1} \lambda_i y_i^2 \geq 0 \qquad (4) \end{aligned}$$

there the inequality in (4) comes from (3). Hence P is positive semi-definite

$$\boxed{\mathbf{x}^*P\mathbf{x} \geq 0}.$$

2.3 Let $\mathbf{A} \in \mathbb{R}^{\mathbf{m \times n}}$ with $\mathbf{m \geq n}$, and let $\mathbf{A} = \hat{\mathbf{Q}}\hat{\mathbf{R}}$ be a reduced QR factorization

The reduced $\hat{Q}\hat{R}$ factorization of A can be written as

$$\begin{pmatrix} | & & | \\ a_1 & \dots & a_n \\ | & & | \end{pmatrix}_{\text{mxn}} = \begin{pmatrix} | & & | \\ q_1 & \dots & q_n \\ | & & | \end{pmatrix}_{\text{mxn}} \begin{pmatrix} \langle q_1, a_1 \rangle & \langle q_1, a_2 \rangle & \dots & \langle q_1, a_n \rangle \\ 0 & \langle q_2, a_2 \rangle & \dots & \langle q_2, a_n \rangle \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \langle q_n, a_n \rangle \end{pmatrix}_{\text{nxn}}$$

Where the columns of $\hat{Q}$ form an orthonormal set of vectors that span the column space of A, and $\hat{R}$ is upper triangular.

2.3.1 Show A has rank n if and only if all the diagonal entries of $\hat{R}$ are nonzero

Consider a vector $\mathbf{v}$ in the nullspace of A. There exists a nonzero solution if and only if all diagonal entries of $\hat{R}$ are nonzero. We can see this easily since $\hat{R}$ is upper triangular and realizing the vectors q_i can only be zero if the same column vector in A, a_i is zero otherwise the product $\hat{Q}\hat{R}$ wouldn't be able to reconstruct the matrix A

2.3.2 Suppose $\hat{R}$ has k nonzero diagonal entries.

This implies the rank of A is exactly k, and furthermore, the columns corresponding to nonzero entries correspond the columns of A, and Q, that form a basis for the column space of A. Consider a basis for column space of A where these vectors can be from any column of A

$$\text{span}(A) = \text{span}(\{a_1, \dots, a_k\}).$$

then preforming Gram Schmidt on this basis form an orthonormal basis for the columns of A, so any other column vector a_{k+1} is in this span

$$a_{k+1} \in \text{span}(\{q_1, \dots, q_k\}).$$

This means that projection of additional vectors, a_{k+1}, onto the othogonal compliment of $\hat{Q}_k\hat{Q}_k$ can be written as a linear combination of these basis vectors where $\hat{Q}_k$ contains the column vectors $\{q_1, \dots, q_k\}$

$$q_{k+1} = (I - \hat{Q}_k\hat{Q}_k)a_{k+1} \in \text{span}(\{q_1, \dots, q_k\}).$$

This means q_{k+1} is not needed to reconstruct the column a_{k+1}. So for $\hat{Q}_{k+1}\hat{r}_{k+1}$, where $\hat{r}_{k+1}$ is the $k+1$ column vector of $\hat{R}$

$$a_{k+1} = \hat{Q}_{k+1}\hat{r}_{k+1} = \sum_{i=1}^{k+1} \langle q_i, a_{k+1} \rangle$$

where a_{k+1} can be written as a linear combination of $\{q_1, \ldots, q_k\}$, so the diagonal component of $\hat{R}$, for the vector $\hat{r}_{k+1}$, $\langle q_{k+1}, a_{k+1}\rangle$ can be written as

$$\begin{aligned}\langle q_{k+1}, a_{k+1}\rangle &= \langle q_{k+1}, c_1 q_1 + \cdots + c_k q_k\rangle \\ &= 0\end{aligned}$$

since q_{k+1} is guaranteed to be orthogonal to the vectors $\{q_1, \ldots, q_k\}$.

This shows that only k nonzero diagonal elements in $\hat{R}$ can exist for a matrix of rank k.

2.4 Determine the eigenvalues, determinant, and singular values of a Householder reflector.

For normalized vector $\|\mathbf{u}\| = 1$, the Householder reflector $H = I - 2\mathbf{u}\mathbf{u}^T$ is the reflection across the plane perpendicular to $\mathbf{u}$, where $P = \mathbf{u}\mathbf{u}^T$ is the orthogonal projection onto the line spanned by $\mathbf{u}$. So $H = I - 2P$ is unitary by (2.1), and hence normal $H^*H = HH^*$ therefore by the spectral theorem is diagonalizable by a unitary matrix.

2.4.1 Eigenvalues

Let $H\mathbf{v} = \lambda\mathbf{v}$, then

$$\begin{aligned}\langle H\mathbf{v}, H\mathbf{v}\rangle = \langle \lambda\mathbf{v}, \lambda\mathbf{v}\rangle &= \underbrace{|\lambda|^2\langle \mathbf{v}, \mathbf{v}\rangle} \\ = \langle \mathbf{v}, H^*Hv\rangle &= \underbrace{\langle \mathbf{v}, \mathbf{v}\rangle}\end{aligned}$$

so $|\lambda|^2 = 1 \Rightarrow \boxed{\lambda = \pm 1}$. Since this is a reflection, all vector lengths will be preserved.

2.4.2 Determinant

Since this is a product of eigenvalues

$$\boxed{\det(H) = \pm 1}.$$

Where the sign depends on the dimension of H.

2.4.3 Singular values

Since H is normal, by the spectral theorem

$$\sqrt{H^*H} = \sqrt{UD^*DU^*} = U|D|U_*.$$

The singular values are the absolute value of the eigenvalues

$$\boxed{\sigma = |\pm 1| = 1}.$$

2.5 Show $\mathbf{cond(A^TA) = (cond(A))^2}$

Let $A = U\Sigma V^T$ be the singular value decomposition of A. The goal is to consider the singular values of A^TA and compare them with A. Then

$$A^TA = V\Sigma^2V^T$$

Which is the singular value decomposition of A^TA, so the singular values are σ_i^2; the singular values of A squared. Hence

$$\text{cond}(A^TA) = (\text{cond}(A))^2$$