Basics in Applied Mathematics – Part I: Numerics

Lecture 5: The QR decomposition and Householder reflections; the singular value decomposition

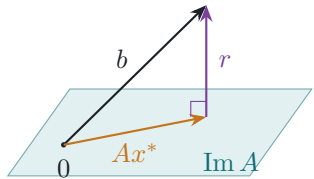
JProf. Dr. Diyora Salimova

Department of Applied Mathematics, University of Freiburg

3 November 2026

Recap of Lecture 4 (1/2): least squares and the normal equations Sections 3.1–3.2

$A \in \mathbb{R}^{m\times n}$, $m > n$ (data fitting: $A = (\phi_j(t_i))_{i,j}$): $Ax = b$ has in general no solution $\Rightarrow$ minimise $\|Ax - b\|_2$.



Theorem 3.4 (normal equations)

x^* is a least squares solution $\iff$ $A^\top A x^* = A^\top b$. A solution always exists; it is unique $\iff$ $\operatorname{rank} A = n$, and then $A^\top A$ is SPD.

Ax^* is the orthogonal projection of b onto $\operatorname{Im} A$; the residual $r = b - Ax^*$ is orthogonal to $\operatorname{Im} A$: $A^\top r = 0$.

Today: the same x^* without forming $A^\top A$ — a QR decomposition and one triangular system (Theorem 3.13).

Recap of Lecture 4 (2/2): the price of the normal equations

Sections 3.2.1, 3.3

- $\operatorname{rank} A = n$: $\operatorname{cond}_2(A) = \sigma_{\max}(A)/\sigma_{\min}(A)$ with $\sigma_{\max}, \sigma_{\min}$ the square roots of the extreme eigenvalues of $A^\top A$ (Definition 3.6); $\operatorname{cond}_2(A^\top A) = \operatorname{cond}_2(A)^2$ (Proposition 3.7)
- Läuchli (Example 3.9): $A = \begin{pmatrix} 1 & 1 \\ \varepsilon & 0 \\ 0 & \varepsilon \end{pmatrix}$, $\varepsilon = 10^{-8}$: $\operatorname{cond}_2(A) \approx 1.4 \cdot 10^8 \ll 1/u$, but the computed $A^\top A$ is singular. **Why?**

$\mathrm{fl}(1 + \varepsilon^2) = 1$, since $\varepsilon^2 = 10^{-16} < u$: the computed $A^\top A$ is $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ — the information is lost before any solver starts.

The way out: orthogonal matrices, $Q^\top Q = I$ (Lemma 3.11)

$Q^{-1} = Q^\top$; $\|Qx\|_2 = \|x\|_2$, $\|QA\|_2 = \|A\|_2$, $\operatorname{cond}_2(Q) = 1$; products of orthogonal matrices are orthogonal.

Today: triangularise A by orthogonal matrices — the QR decomposition (Theorem 3.12), computed with Householder reflections (Section 3.5); then Chapter 4: the SVD.

Theorem 3.12 (QR decomposition)

Let $A \in \mathbb{R}^{m\times n}$ with $m \geq n$. There are an orthogonal $Q \in \mathbb{R}^{m\times m}$ and an upper triangular $R_1 \in \mathbb{R}^{n\times n}$ with

$$A = Q \begin{pmatrix} R_1 \\ 0 \end{pmatrix} = \begin{pmatrix} Q_1 & Q_2 \end{pmatrix} \begin{pmatrix} R_1 \\ 0 \end{pmatrix} = Q_1 R_1,$$

where $Q_1 \in \mathbb{R}^{m\times n}$ holds the first n columns of Q and $Q_2 \in \mathbb{R}^{m\times(m-n)}$ the rest. If $\operatorname{rank} A = n$, then R_1 is regular, and Q_1, R_1 are unique once the diagonal entries of R_1 are required to be positive.

- $A = Q_1 R_1$: the **reduced** QR decomposition;
- regularity: $n = \operatorname{rank}(Q_1 R_1) \leq \operatorname{rank} R_1 \leq n$; uniqueness: in the notes
- existence: **constructively**, by Householder reflections (Section 3.5, Proposition 3.23)

Theorem 3.13 (least squares via QR)

Let $\operatorname{rank} A = n$ and $A = Q\begin{pmatrix} R_1 \\ 0 \end{pmatrix} = Q_1 R_1$ as in Theorem 3.12. Then the least squares solution of $Ax = b$ is the solution of the regular triangular system

$$R_1 x^* = Q_1^\top b, \qquad \text{and} \qquad \min_{x \in \mathbb{R}^n} \|Ax - b\|_2 = \|Q_2^\top b\|_2.$$

Blackboard: proof.

Remark 3.14 (no loss of conditioning). $A^\top A = R_1^\top Q_1^\top Q_1 R_1 = R_1^\top R_1 \;\Rightarrow\; A$ and R_1 have the same $\sigma_{\max}, \sigma_{\min}$, so

$$\operatorname{cond}_2(R_1) = \operatorname{cond}_2(A) \qquad \text{instead of} \qquad \operatorname{cond}_2(A^\top A) = \operatorname{cond}_2(A)^2.$$

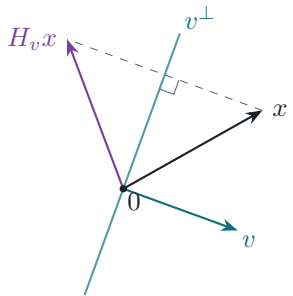
The matrix of the triangular system has the condition number of A, not its square.

Householder reflections

Idea: *triangularise* A by orthogonal transformations — elimination with reflections in place of the E_k (Lemma 2.11).

Definition 3.19. For $v \in \mathbb{R}^m \setminus \{0\}$ the **Householder reflection** is

$$H_v := I - 2\,\frac{vv^\top}{v^\top v} \in \mathbb{R}^{m\times m}.$$



mirror image of x in $v^\perp$

Lemma 3.20 (properties)

(i) H_v is symmetric and orthogonal, and $H_v^2 = I$;

(ii) $H_v v = -v$ and $H_v w = w$ for $w \perp v$: H_v is the **reflection** in the hyperplane $v^\perp$;

(iii) $H_v x = x - 2\,\frac{v^\top x}{v^\top v}\,v$ costs $O(m)$ flops — the matrix H_v is never formed.

Choosing the reflection

Lemma 3.21 (choice of the reflection)

$x \in \mathbb{R}^m \setminus \{0\}$, $\operatorname{sgn}(x_1) \in \{\pm 1\}$ with $\operatorname{sgn} 0 := 1$, $\alpha := -\operatorname{sgn}(x_1)\,\|x\|_2$, $v := x - \alpha e_1$. Then $v \neq 0$ and $H_v x = \alpha e_1 = -\operatorname{sgn}(x_1)\,\|x\|_2\, e_1$.

Blackboard: proof.

Why this sign: $v_1 = x_1 + \operatorname{sgn}(x_1)\|x\|_2$ adds numbers of the same sign — *no cancellation*

Example: $x = (3,4)^\top$: $\|x\|_2 = 5$, $\alpha = -5$, $v = (8,4)^\top$,

$$H_v = I - \frac{2\,vv^\top}{80} = \frac{1}{5}\begin{pmatrix} -3 & -4 \\ -4 & 3 \end{pmatrix}, \qquad H_v x = \begin{pmatrix} -5 \\ 0 \end{pmatrix} = \alpha e_1.$$

Algorithm 3.22 (Householder QR decomposition)

$A \in \mathbb{R}^{m\times n}$, $m \geq n$, $A^{(1)} := A$. For $k = 1, \ldots, p := \min(n, m-1)$:
$x^{(k)} := \left(a_{kk}^{(k)}, \ldots, a_{mk}^{(k)}\right)^\top \in \mathbb{R}^{m-k+1}$;
if $x^{(k)} = 0$: $H_k := I$; otherwise v_k for $x^{(k)}$ by Lemma 3.21 and

$$H_k := \begin{pmatrix} I_{k-1} & 0 \\ 0 & H_{v_k} \end{pmatrix}, \qquad A^{(k+1)} := H_k A^{(k)},$$

that is, H_{v_k} is applied to rows $k, \ldots, m$ of columns $k, \ldots, n$.

$m = 4$, $n = 3$ ($*$: final entries, $\times$: still to be changed):

$$\begin{pmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{pmatrix} \xrightarrow{H_1} \begin{pmatrix} * & * & * \\ 0 & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \end{pmatrix} \xrightarrow{H_2} \begin{pmatrix} * & * & * \\ 0 & * & * \\ 0 & 0 & \times \\ 0 & 0 & \times \end{pmatrix} \xrightarrow{H_3} \begin{pmatrix} * & * & * \\ 0 & * & * \\ 0 & 0 & * \\ 0 & 0 & 0 \end{pmatrix}$$

Householder QR gives Theorem 3.12 — and solves least squares problems

Proposition 3.23

In Algorithm 3.22 every H_k is symmetric and orthogonal, step k leaves rows $1, \dots, k-1$ unchanged, and $A^{(k+1)}$ has zeros below the diagonal in columns $1, \dots, k$. Hence

$$H_p \cdots H_1 A = \begin{pmatrix} R_1 \\ 0 \end{pmatrix}, \qquad Q := H_1 \cdots H_p \text{ orthogonal}, \qquad A = Q \begin{pmatrix} R_1 \\ 0 \end{pmatrix}:$$

Theorem 3.12 holds for every $A \in \mathbb{R}^{m \times n}$ with $m \geq n$.

Order of the factors: $Q = H_1 \cdots H_p$, but $Q^\top = H_p \cdots H_1$ (the H_k are symmetric) — applying H_1 first and H_p last to a vector y gives $Q^\top y$. Proof: induction on k (notes).
$\operatorname{rank} A = n$: **do not form Q.**

- apply $H_1, \dots, H_p$ to b as well (in this order) $\Rightarrow$ $Q^\top b$
- solve the regular triangular system $R_1 x = (Q^\top b)_{1:n}$ (backward substitution)
- the remaining components give the minimal residual: $\|(Q^\top b)_{n+1:m}\|_2 = \|Q_2^\top b\|_2$ (Theorem 3.13)

Example 3.25: Householder QR by hand

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -\sqrt{2} & \sqrt{2}/2 \\ 0 & \sqrt{2} & 5/\sqrt{2} \end{pmatrix}, \qquad b = \begin{pmatrix} 3\sqrt{2} \\ -1 \\ 7 \end{pmatrix}.$$

Blackboard: two reflection steps, $Q^\top b$, backward substitution.

$$R = \begin{pmatrix} -1 & -1 & -1 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix}, \qquad Q^\top b = \begin{pmatrix} -3\sqrt{2} \\ 4\sqrt{2} \\ 3\sqrt{2} \end{pmatrix}, \qquad x = \begin{pmatrix} \sqrt{2} \\ \sqrt{2} \\ \sqrt{2} \end{pmatrix} \qquad \text{(check: } Ax = b\text{)}.$$

$r_{11} = -1 < 0$: Householder's sign choice does not give the positive diagonal of Theorem 3.12 — flip the signs of row 1 of R and column 1 of Q.
Computed (`householder_qr` in the notes): $\|A - QR\|_F = 5 \cdot 10^{-16}$, $\|Q^\top Q - I\|_F = 1 \cdot 10^{-17}$.

Summary of Chapter 3

- **least squares**: minimise $\|Ax - b\|_2$; solutions $\iff$ normal equations $A^\top Ax = A^\top b$; the residual is orthogonal to $\operatorname{Im} A$; unique $\iff \ker A = \{0\}$
- $\operatorname{rank} A = n$: solving the normal equations **squares the condition number**, $\operatorname{cond}_2(A^\top A) = \operatorname{cond}_2(A)^2$ — information is lost when $A^\top A$ is formed
- **QR** ($\operatorname{rank} A = n$): $A = Q_1 R_1$ reduces the problem to the regular triangular system $R_1 x = Q_1^\top b$ with $\operatorname{cond}_2(R_1) = \operatorname{cond}_2(A)$; the minimal residual is $\|Q_2^\top b\|_2$
- **Householder**: $H_v = I - 2\,vv^\top/(v^\top v)$, $v = x - \alpha e_1$, $\alpha = -\operatorname{sgn}(x_1)\|x\|_2$

QR: one orthogonal matrix *triangularises* A. With orthogonal matrices on *both* sides every rectangular matrix can be *diagonalised*:

$$A = U\Sigma V^\top.$$

Used for rank-deficient least squares, numerical rank and cond_2, image compression, principal component analysis.

After this chapter you can

(i) state and prove the existence of the singular value decomposition (SVD),
(ii) read rank, norm, condition number, image and kernel of a matrix off its SVD,
(iii) use the pseudoinverse to solve rank-deficient least squares problems,
(iv) understand and apply the best low-rank approximation theorem (image compression).

Today (i) and (ii): 4.1 The theorem · 4.2 What the SVD tells us about a matrix

Spectral theorem: $S \in \mathbb{R}^{n\times n}$ symmetric $\Rightarrow$ $S = V\Lambda V^\top$, V orthogonal, $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ real; S positive semidefinite $\Rightarrow$ $\lambda_i = v_i^\top S v_i \geq 0$.

Theorem 4.1 (singular value decomposition)

Let $A \in \mathbb{R}^{m\times n}$ and $p := \min(m, n)$. There exist orthogonal matrices $U \in \mathbb{R}^{m\times m}$, $V \in \mathbb{R}^{n\times n}$ and $\Sigma \in \mathbb{R}^{m\times n}$ with $\Sigma_{ij} = 0$ for $i \neq j$, $\Sigma_{ii} = \sigma_i$, $\sigma_1 \geq \cdots \geq \sigma_p \geq 0$, such that

$$A = U\Sigma V^\top.$$

The **singular values** $\sigma_1, \ldots, \sigma_p$ are uniquely determined by A. The columns u_i of U and v_i of V (left and right **singular vectors**) satisfy $Av_i = \sigma_i u_i$ and $A^\top u_i = \sigma_i v_i$, $i = 1, \ldots, p$.

Shape of Σ: $m = 3$, $n = 2$: $\begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{pmatrix}$, $m = 2$, $n = 3$: $\begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \end{pmatrix}$.

Reduced SVD and geometry

$r :=$ number of nonzero singular values. **Reduced SVD** (first r columns of U and V):

$$A = U_r \Sigma_r V_r^\top = \sum_{i=1}^{r} \sigma_i \, u_i v_i^\top, \qquad U_r \in \mathbb{R}^{m\times r}, \;\; \Sigma_r = \operatorname{diag}(\sigma_1, \ldots, \sigma_r), \;\; V_r \in \mathbb{R}^{n\times r}, \quad (4.1)$$

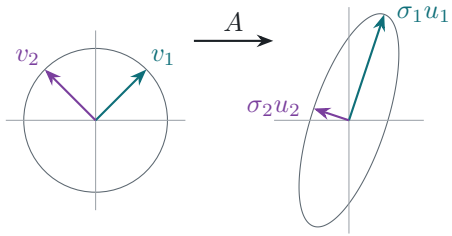
a sum of r rank-one matrices with decreasing weights.

Geometry: A maps the orthonormal basis $v_1, \ldots, v_n$ to the orthogonal vectors $\sigma_1 u_1, \ldots, \sigma_n u_n$ ($\sigma_i := 0$ for $i > p$): the unit sphere becomes an **ellipsoid** with semi-axes $\sigma_i u_i$.

For instance $A = \begin{pmatrix} 3 & 0 \\ 4 & 5 \end{pmatrix}$: $\sigma_1 = 3\sqrt{5}$, $\sigma_2 = \sqrt{5}$,

$v_1 = \frac{1}{\sqrt{2}}(1,1)^\top$, $v_2 = \frac{1}{\sqrt{2}}(-1,1)^\top$,

$\sigma_1 u_1 = \frac{1}{\sqrt{2}}(3,9)^\top$, $\sigma_2 u_2 = \frac{1}{\sqrt{2}}(-3,1)^\top$.



unit circle $\mapsto$ ellipse

Example 4.2: singular values by hand

$A = \begin{pmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 3}$. The nonzero eigenvalues of $AA^\top$ and $A^\top A$ coincide (proof of Proposition 1.23(i)) — use the smaller matrix.

Recipe: $AA^\top = ?$ its eigenvalues ? $\sigma_1, \sigma_2 = ?$

$AA^\top = \begin{pmatrix} 11 & 1 \\ 1 & 11 \end{pmatrix}$, eigenvalues 12 and 10 $\Rightarrow$ $\sigma_1 = \sqrt{12} \approx 3.46$, $\sigma_2 = \sqrt{10} \approx 3.16$;
$\operatorname{rank} A = 2$, and $A^\top A \in \mathbb{R}^{3\times 3}$ has the eigenvalues 12, 10, 0.

Proposition 4.3

Let $A = U\Sigma V^\top \in \mathbb{R}^{m\times n}$ with $\sigma_1 \geq \cdots \geq \sigma_r > 0 = \sigma_{r+1} = \cdots = \sigma_p$. Then:

(i) $\operatorname{rank} A = r$, $\operatorname{Im} A = \operatorname{span}(u_1, \ldots, u_r)$ and $\ker A = \operatorname{span}(v_{r+1}, \ldots, v_n)$;

(ii) $\|A\|_2 = \sigma_1$ and $\|A\|_F = \left(\sigma_1^2 + \cdots + \sigma_r^2\right)^{1/2}$;

(iii) $m = n$, A regular: $\|A^{-1}\|_2 = 1/\sigma_n$ and $\operatorname{cond}_2(A) = \sigma_1/\sigma_n$;

(iv) $m = n$, A symmetric with eigenvalues $\lambda_1, \ldots, \lambda_n$: the singular values are $|\lambda_1|, \ldots, |\lambda_n|$ (in decreasing order);

(v) $m = n$: $\sigma_n = \min\{\|A - B\|_2 : B \in \mathbb{R}^{n\times n} \text{ singular}\}$, the distance of A to the singular matrices.

Summary

- **SVD**: every $A \in \mathbb{R}^{m\times n}$ is $A = U\Sigma V^\top$ with U, V orthogonal and Σ diagonal, $\sigma_1 \geq \cdots \geq \sigma_p \geq 0$; $\sigma_1^2, \ldots, \sigma_p^2$ are the p largest eigenvalues of $A^\top A$; $Av_i = \sigma_i u_i$
- reduced SVD $A = \sum_{i=1}^{r} \sigma_i u_i v_i^\top$: A maps the unit sphere to an ellipsoid with semi-axes $\sigma_i u_i$
- read off the SVD: $\operatorname{rank} A = r$, $\operatorname{Im} A$, $\ker A$, $\|A\|_2 = \sigma_1$, $\|A\|_F$, $\operatorname{cond}_2(A) = \sigma_1/\sigma_n$, the distance σ_n to the singular matrices