Lecture 22 - Matrix Decompositions: QR and LU

November 18, 2022

Goals: Be able to solve linear systems / compute the inverse of a matrix given a suitable decomposition.

Recall that a matrix A is **upper triangular** if all entries "below the main diagonal" are 0, i.e., $a_{ij} = 0$ for $j < i$ (the diagonal can have 0s). Similarly, A is **lower triangular** if all entries "above the main diagonal" are 0, i.e. $a_{ij} = 0$ for $i < j$.

If A is an upper triangular square matrix with nonzero diagonal entries. We have a systematic way of solving the system $A\mathbf{x} = \mathbf{b}$ for any vector $\mathbf{b}$.

Example 1: (Back-substitution) Solve the system equations:

$$\begin{aligned} x_1 + 3x_2 - 2x_3 &= 9 \\ x_2 - x_3 &= 3 \\ 2x_3 &= 4 \end{aligned}$$

$\rightarrow A = \begin{bmatrix} 1 & 3 & -2 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{bmatrix}$

$x_1 + 15 - 4 = 9 \rightarrow x_1 = -2$

$x_2 - 2 = 3 \rightarrow x_2 = 5$

$x_3 = 2$

$\vec{x} = \begin{bmatrix} -2 \\ 5 \\ 2 \end{bmatrix}$

assuming $A\vec{x} = \vec{b}$, since there's a unique solution, $\vec{x} = A^{-1}\vec{b}$ and A is invertible

Similarly, if A is a lower triangular matrix with nonzero diagonal entries, we can perform the same procedure starting with x_1 instead of x_n (we will still call this back-substitution). This explains the following result:

Theorem 22.1.3: If A is an upper or lower triangular $n \times n$ matrix with all diagonal entries nonzero, then A is invertible.

Conversely, an upper/lower triangular matrix with at least one 0 entry on the diagonal is **never** invertible (it's null space would be nonzero).

Let us now consider under/over determined systems involving triangular matrices.

Example 2: (Underdetermined upper triangular) Solve

$n > m$, less equations than # of unknowns

$$\begin{aligned} 2x_1 - 3x_2 + x_3 + 3x_4 - x_5 &= 2 \\ x_2 + 4x_3 + 4x_4 - 7x_5 &= 5 \\ 3x_3 + 3x_4 - 6x_5 &= 3 \end{aligned}$$

$m \times n \rightarrow 3 \times 5$

$A = \begin{bmatrix} 2 & -3 & 1 & 3 & -1 \\ 0 & 1 & 4 & 4 & -7 \\ 0 & 0 & 3 & 3 & -6 \end{bmatrix}$

We move equations we don't know to other side, then back solve like Ex 1

$$\left.\begin{aligned} 2x_1 - 3x_2 + x_3 &= 2 - 3x_4 + x_5 \\ x_2 + 4x_3 &= 5 - 4x_4 + 7x_5 \\ 3x_3 &= 3 - 3x_4 + 6x_5 \end{aligned}\right\} \rightarrow \begin{aligned} x_3 &= 1 - x_4 + 2x_5 \\ x_2 &= -x_5 + 1 \\ x_1 &= 2 - x_4 + 2x_5 \end{aligned}$$

$\dim C(A) + \dim N(A) = n = 5$

Solutions: $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 2 - x_4 + 2x_5 \\ -x_5 + 1 \\ 1 - x_4 + 2x_5 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -1 \\ 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} + x_5 \begin{pmatrix} 2 \\ -1 \\ 2 \\ 0 \\ 1 \end{pmatrix}$

$\Rightarrow$ infinitely many solutions

$C(A) = \mathbb{R}^3$ ($m = 3$) bc 3 vectors

$\therefore \dim N(A) = 2$ $\left(N(A) = \text{span}\left\{\begin{pmatrix} -1 \\ 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ 2 \\ 0 \\ 1 \end{pmatrix}\right\}\right)$

Example 3: (Underdetermined lower triangular) Solve

$$\begin{bmatrix}2&0&0&0&0\\-3&1&0&0&0\\1&4&3&0&0\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{bmatrix}=\begin{bmatrix}6\\-7\\5\end{bmatrix}$$

$$C(A) = \text{span}\left(\begin{pmatrix}2\\-3\\1\end{pmatrix}, \begin{pmatrix}0\\1\\4\end{pmatrix}, \begin{pmatrix}0\\0\\3\end{pmatrix}, \overbrace{\begin{pmatrix}0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\end{pmatrix}}^{\text{drop these}}\right)$$

$$2x_1 = 6$$
$$-3x_1 + x_2 = -7$$
$$x_1 + 4x_2 + 3x_3 = 5$$

$$\Rightarrow \vec{x} = \begin{pmatrix}3\\2\\-2\\x_4\\x_5\end{pmatrix}$$

solutions: $\vec{x} = \begin{pmatrix}3\\2\\-2\\0\\0\end{pmatrix} + x_4\begin{pmatrix}0\\0\\6\\1\\0\end{pmatrix} + x_5\begin{pmatrix}0\\0\\6\\0\\1\end{pmatrix}$

infinitely many solutions

$\underbrace{C(A) = \mathbb{R}^3}_{3 \text{ 3-vectors}}$, $\dim N(A) = 2$

Example 4: (Overdetermined upper triangular) Solve

$m > n$, more eqns than unknowns

$$\begin{array}{c}5\times 3\\ \begin{bmatrix}2&-3&1\\0&1&4\\0&0&3\\0&0&0\\0&0&0\end{bmatrix}\end{array}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}b_1\\b_2\\b_3\\b_4\\b_5\end{bmatrix}$$

these have unique solution

$$2x_1 - 3x_2 + x_3 = b_1$$
$$x_2 + 4x_3 = b_2$$
$$3x_3 = b_3$$

g. condition

$$0 = b_4$$
$$0 = b_5$$

back-sub $\Rightarrow$ $\vec{x} = \begin{pmatrix}\frac{1}{2}b_1 + \frac{3}{2}b_2 - \frac{13}{6}b_3\\ b_2 - \frac{4}{3}b_3\\ \frac{1}{3}b_3\end{pmatrix}$

unique solution iff $b_4 = 0$ & $b_5 = 0$

$\dim N(A) = 0$

$\underbrace{\dim C(A) = 3}_{3 \text{ 5-vectors}}$, $C(A) \subset \mathbb{R}^5$

Example 5: (Overdetermined lower triangular) Solve

$$\begin{bmatrix}2&0&0\\-3&1&0\\1&-4&3\\2&-7&3\\4&5&-1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}b_1\\b_2\\b_3\\b_4\\b_5\end{bmatrix}$$

$\hookrightarrow$ same idea as before

$\dim C(A) = 3$

$\dim N(A) = 0$

To summarize:

Theorem 22.1.9: Let A be an upper or lower triangular $m \times n$ matrix with all diagonal entries nonzero.

1. If $m = n$, then $A\mathbf{x} = \mathbf{b}$ always has a unique solution (via back-substitution).
   square
2. In the underdetermined case $(m < n)$, $A\mathbf{x} = \mathbf{b}$ always has a solution (so $C(A) = \mathbb{R}^m$) and $\dim N(A) = n - m > 0$. The solutions are given by uniquely solving for $x_1, \dots x_m$ in terms of the arbitrarily chosen $x_{m+1}, \dots, x_n$. For lower triangular A, the system $A\mathbf{x} = \mathbf{b}$ is really m equations in $x_1, \dots, x_m$ with the other x_j's having coefficient 0 in every equation.
   less eqns, more unknowns
3. In the overdetermined case $(m > n)$, $\dim C(A) = n < m$ (so a solution exists only for special $\mathbf{b}$) and $N(A) = \{\mathbf{0}\}$ (so a solution to $A\mathbf{x} = \mathbf{b}$ is unique when one exists). The column space $C(A)$ is described by expressions for each of $b_{n+1}, \dots, b_m$ in terms of $b_1, \dots, b_n$; for upper triangular A these expressions are simply $b_{n+1} = \dots = b_m = 0$.
   more eqns, less unknowns

Before we continue, we more formally state a result we saw at the end of lecture 20:

Theorem 22.1.4: For an $n \times n$ orthogonal matrix A, the equation $A\mathbf{x} = \mathbf{b}$ has the unique solution $\mathbf{x} = A^{\mathsf{T}}\mathbf{b}$.

$A^{-1} = A^T$ for orthog

Matrix Decompositions

Theorem 22.2.1:

1. (LU-decomposition) "Most" $n \times n$ matrices A have the form $A = LU$ for $n \times n$ lower triangular L and $n \times n$ upper triangular U. The matrix A is invertible precisely when the diagonal entries of L and U are nonzero.
2. (QR-decomposition) An invertible $n \times n$ matrix A can be written as $A = QR$ where Q is an $n \times n$ orthogonal matrix and R is an $n \times n$ upper triangular matrix with **positive** diagonal entries.

Example 6: Suppose that we are given the following LU-decomposition for A:

$$A = \begin{bmatrix} 2 & -2 & -3 \\ 4 & -3 & -4 \\ -6 & 6 & 14 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 2 & -1 & 0 \\ -3 & 0 & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} 2 & -2 & -3 \\ 0 & -1 & -2 \\ 0 & 0 & 5 \end{bmatrix}}_{U}.$$

Solve the following system of linear equations:

$$\begin{aligned} 2x - 2y \ -3z &= 0 \\ 4x - 3y \ -4z &= 6 \\ -6x + 6y + 14z &= 10 \end{aligned}$$

$A\vec{x} = \vec{b} \to L(U\vec{x}) = \vec{b} \to$ suppose $U\vec{x} = \vec{y}$, (1) solve $L\vec{y} = \vec{b}$ (2) solve $U\vec{x} = \vec{y}$ } change it into two easy systems

(1) $L\vec{y} = \vec{b}$

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & -1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 6 \\ 10 \end{bmatrix} \to \vec{y} = \begin{bmatrix} 0 \\ -6 \\ 10 \end{bmatrix}$$

(2) solve $U\vec{x} = \vec{y}$

$$\begin{bmatrix} 2 & -2 & -3 \\ 0 & -1 & -2 \\ 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ -6 \\ 10 \end{bmatrix} \to \boxed{\vec{x} = \begin{bmatrix} 5 \\ 2 \\ 2 \end{bmatrix}}$$

Example 7: How would you solve the system $A\mathbf{x} = \mathbf{b}$ given the QR decomposition of A?

$$A\vec{x} = \vec{b} \to Q(R\vec{x}) = \vec{b} \to Q^TQ(R\vec{x}) = Q^T\vec{b} \to \underbrace{R\vec{x} = Q^T\vec{b}}$$

since Q is orthog, $Q^{-1} = Q^T$

multiply by transpose

solve via back-sub

Computing Inverses of Triangular Matrices

If we have the LU or the QR decomposition of a matrix A, then we can compute A^{-1} as follows:

- $A = LU$ implies $A^{-1} = U^{-1}L^{-1}$.
- $A = QR$ implies $A^{-1} = R^{-1}Q^{\top}$.

Thus, if we know how to invert upper and lower triangular matrices, we can invert A.

Fact: The inverse of an upper triangular is upper triangular, and the inverse of a lower triangular is lower triangular.

Example 8: Find the inverse of the following matrices:

$$\overset{U}{A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}}, \quad \overset{L}{B = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 3 & -1 \end{bmatrix}}$$

$$AA^{-1} = I_n$$

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} d_1 & a & b \\ 0 & d_2 & c \\ 0 & 0 & d_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$d_1 + 0 + 0 = 1 \to d_1 = 1$

$a + d_2 + 0 = 0 \quad a + \frac{1}{2} = 0 \to a = -\frac{1}{2}$

$b + c + 0 = 0 \quad b - \frac{1}{2} = 0 \to b = \frac{1}{2}$

$0 + 0 + 0 = 0$

$0 + 2d_2 + 0 = 1 \to d_2 = \frac{1}{2}$

$0 + 2c + d_3 = 0 \to 2c + 1 = 0 \to c = -\frac{1}{2}$

$0 + 0 + 0 = 0$

$0 + 0 + 0 = 0$

$0 + 0 + d_3 = 1 \to d_3 = 1$

diagonal entries are reciprocals

$$\to A^{-1} = \begin{bmatrix} 1 & -1/2 & 1/2 \\ 0 & 1/2 & -1/2 \\ 0 & 0 & 1 \end{bmatrix}$$

diagonals of inverse are always reciprocal

Given an $n \times n$ matrix A with linearly independent columns, we can compute its QR decomposition from the Gram-Schmidt process. We will illustrate the process through the following example (see page 495 of the text for the explicit steps).

Example 9: Determine the QR decomposition of the matrix

$$A = \begin{bmatrix} -2 & 1 \\ 1 & 3 \end{bmatrix}.$$

to do this, must use first column as v_1

Gram Schmidt! $\vec{v}_1 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$

$\vec{w}_1 = \vec{v}_1 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$

$\vec{w}_2 = \vec{v}_2 - \text{Proj}_{\vec{w}_1}(\vec{v}_2) = \begin{bmatrix} 7/5 \\ 14/5 \end{bmatrix}$

$\vec{w}_1{}' = \frac{\vec{w}_1}{\|\vec{w}_1\|} = \frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 1 \end{bmatrix}$

$\vec{w}_2{}' = \frac{\vec{w}_2}{\|\vec{w}_2\|} = \frac{\sqrt{5}}{7} \begin{bmatrix} 7/5 \\ 14/5 \end{bmatrix}$

$$Q = \begin{bmatrix} | & | \\ \vec{w}_1{}' & \vec{w}_2{}' \\ | & | \end{bmatrix}$$

orthog matrix, covers same span (v_1, v_2) b/c it's a basis

To get R, write $\vec{v}_i$ as a linear combination of the $\vec{w}_i{}'$. R is coefficients.

↳ gives us upper triangular b/c the process of Gram-Schmidt gives us vectors that are orthogonal to others

(1) $\vec{v}_1 = \vec{w}_1 = \sqrt{5}\, \vec{w}_1{}'$

(2) $\vec{w}_2 = \vec{v}_2 - \text{Proj}_{\vec{w}_1}(\vec{v}_2)$

$\vec{v}_2 = \text{Proj}_{\vec{w}_1}(\vec{v}_2) + \vec{w}_2 = \frac{1}{5}\vec{w}_1 + \vec{w}_2 = \frac{\sqrt{5}}{5}\vec{w}_1{}' + \frac{7}{\sqrt{5}}\vec{w}_2{}'$

$$R = \begin{bmatrix} \sqrt{5} & \sqrt{5}/5 \\ 0 & 7/\sqrt{5} \end{bmatrix}$$

We shall make a small but important refinement in the Gram–Schmidt process that was absent in our discussion in Chapter 19: at the end we divide each of the (nonzero) "output" vectors $\mathbf{w}_i$ by its length to obtain a set of vectors that is orthonormal (not just mutually orthogonal). To avoid confusion with the notation used in our discussion of the Gram–Schmidt process in Chapter 19, we denote these new unit vectors as $\mathbf{w}'_1, \mathbf{w}'_2, \mathbf{w}'_3$ (so $\mathbf{w}'_i = \mathbf{w}_i/\|\mathbf{w}_i\|$).

Let us see how these unit vectors $\mathbf{w}'_i$ are related to the $\mathbf{v}_j$'s. $\vec{v}_1 = \|\vec{v}_1\| \vec{w}_1'$

(i) In the first step, we have $\mathbf{w}_1 = \mathbf{v}_1$, so $\mathbf{w}'_1 = \mathbf{w}_1/\|\mathbf{w}_1\| = \mathbf{v}_1/\|\mathbf{v}_1\|$. In other words, $\mathbf{v}_1 = r_{11}\mathbf{w}'_1$ where r_{11} is the scalar $\|\mathbf{v}_1\| > 0$.

(ii) At the next step, the orthogonal pair of unit vectors $\mathbf{w}'_1, \mathbf{w}'_2$ has the same span as $\mathbf{w}_1, \mathbf{w}_2$, which has the same span as $\mathbf{v}_1, \mathbf{v}_2$, and by design $\mathbf{w}_2 = \mathbf{v}_2 - t\mathbf{w}_1$ for some scalar t. Hence, $\mathbf{v}_2 = t\mathbf{w}_1 + \mathbf{w}_2 = r_{12}\mathbf{w}'_1 + r_{22}\mathbf{w}'_2$ for some scalars r_{12} and r_{22}, with $r_{22} = \|\mathbf{w}_2\| > 0$.

(iii) Similarly, the orthonormal triple $\mathbf{w}'_1, \mathbf{w}'_2, \mathbf{w}'_3$ has the same span as $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, with $\mathbf{w}_3 = \mathbf{v}_3 - a\mathbf{v}_1 - b\mathbf{v}_2$ for some scalars a, b, so $\mathbf{v}_3 = a\mathbf{v}_1 + b\mathbf{v}_2 + \mathbf{w}_3 = r_{13}\mathbf{w}'_1 + r_{23}\mathbf{w}'_2 + r_{33}\mathbf{w}'_3$ for scalars r_{13}, r_{23}, r_{33} with $r_{33} = \|\mathbf{w}_3\| > 0$.

The Gram–Schmidt process gives formulas for the scalar coefficients $(r_{11}, r_{12}, r_{22}, r_{13}, r_{23}, r_{33})$ in terms of the dot products $\mathbf{v}_i \cdot \mathbf{v}_j$; we illustrate this explicitly in Example 22.4.3 below.

Now we write this in matrix language and see that it expresses exactly the QR-decomposition for A. Let Q be the 3×3 matrix whose columns are $\mathbf{w}'_1, \mathbf{w}'_2, \mathbf{w}'_3$ in turn; i.e.,

$$Q = \begin{bmatrix} | & | & | \\ \mathbf{w}'_1 & \mathbf{w}'_2 & \mathbf{w}'_3 \\ | & | & | \end{bmatrix}.$$

By design, the collection of columns of Q is orthonormal. The three equalities in (i), (ii), (iii) above correspond to the three equalities

$$Q\begin{bmatrix} r_{11} \\ 0 \\ 0 \end{bmatrix} = \mathbf{v}_1, \quad Q\begin{bmatrix} r_{12} \\ r_{22} \\ 0 \end{bmatrix} = \mathbf{v}_2, \quad Q\begin{bmatrix} r_{13} \\ r_{23} \\ r_{33} \end{bmatrix} = \mathbf{v}_3,$$

which when put together amounts to the matrix equation

$$Q\begin{bmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{bmatrix} = \begin{bmatrix} | & | & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \\ | & | & | \end{bmatrix} = A.$$

But Q has *orthonormal* columns, so this is exactly the promised QR-decomposition of the matrix A (since the diagonal entries r_{11}, r_{22}, r_{33} of R are positive).

$$A = \begin{bmatrix} | & | & | \\ v_1 & v_2 & v_3 \\ | & | & | \end{bmatrix} \qquad Q = \begin{bmatrix} | & | & | \\ \frac{w_1'}{\|w_1\|} & \frac{w_2'}{\|w_2\|} & \frac{w_3'}{\|w_3\|} \\ | & | & | \end{bmatrix}$$

Gram Schmidt:

$w_1 = v_1$, $\quad w_1' = \dfrac{w_1}{\|w_1\|}$

$v_1 = \|w_1\| \, w_1'$

$$R = \begin{array}{c} \begin{array}{ccc} v_1 & v_2 & v_3 \\ \downarrow & \downarrow & \downarrow \end{array} \\ \begin{bmatrix} \|w_1\| & c\|w_1\| & \\ 0 & \|w_2\| & \\ 0 & 0 & \end{bmatrix} \end{array}$$

each column i is weights of $\vec{w}_j'$ for v_i

$$w_2 = v_2 - \mathrm{Proj}_{w_1}(\vec{v}_2) = v_2 - \frac{w_1 \cdot v_2}{w_1 \cdot w_1} w_1 = v_2 - c w_1$$

$$v_2 = w_2 + c w_1 = \|w_2\| w_2' + c\|w_1\| w_1'$$