

7.8 Special Determinants

A determinant of size n has $n!$ terms, so a general formula for a large matrix is hopeless. Some matrices, though, have so much structure that their determinants collapse to a short product or a recurrence. This section computes the most useful of them: the Vandermonde determinant, which governs polynomial interpolation; tridiagonal determinants, which satisfy a three-term recurrence; block triangular determinants, which split into smaller ones; and the determinant of a rank-one update of the identity. The proofs share one idea: find the structure first (a polynomial with known roots, a row with two nonzero entries, a pattern of zeros) and let the defining properties of the determinant do the rest.

7.8.1 The Vandermonde determinant

Chapter 3 proved that there is exactly one polynomial of degree at most n through $n+1$ points with distinct first coordinates (Theorem 3.7.3). Behind it sat a square matrix of powers of the nodes. Its determinant turns out to be a product of all the differences of the nodes, and that product explains at a glance why distinct nodes are exactly what is needed.

For $x_1, \dots, x_n \in F$, the **Vandermonde matrix** is

$$\mathbf{V}(x_1, \dots, x_n) := \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{pmatrix} \in M_n(F),$$

whose (i, j) -entry is x_i^{j-1} . Row i lists the powers $1, x_i, \dots, x_i^{n-1}$ of the i -th node. For $n = 1$ it is the 1×1 matrix (1) , and for $n = 2$ its determinant is $x_2 - x_1$.

Theorem 7.8.1. Vandermonde Determinant

Let $n \geq 1$ and $x_1, \dots, x_n \in F$. Then

$$\det \mathbf{V}(x_1, \dots, x_n) = \prod_{1 \leq i < j \leq n} (x_j - x_i),$$

where the empty product (for $n = 1$) is 1.

Idea. Induction on n . Replace the last node x_n by a variable t . Expanding along the last row shows the determinant is a polynomial $f(t)$ of degree at most $n-1$, with leading coefficient the smaller Vandermonde determinant. Setting $t = x_i$ for $i < n$ makes two rows equal, so $x_1, \dots, x_{n-1}$ are roots of f . A polynomial of degree at most $n-1$ with these $n-1$ roots is its leading coefficient times $\prod_{i < n} (t - x_i)$. Now put $t = x_n$ back. The roots argument needs the x_i distinct; if they are not, both sides are 0 anyway.

Proof. We use induction on n (Theorem 0.2.7). For $n = 1$, $\det(1) = 1$, the empty product.

Let $n \geq 2$, assume the formula for $n-1$ nodes, and write $\mathbf{V} = \mathbf{V}(x_1, \dots, x_n)$. Let $C_{n1}, \dots, C_{nn}$ be the cofactors of $\mathbf{V}$ along row n (Definition 7.6.1). Each C_{nj} is computed after deleting

row n , so it does not involve x_n . Define

$$f(t) := \sum_{j=1}^n C_{nj} t^{j-1} \in F[t].$$

For $c \in F$, let $\mathbf{V}_c$ be $\mathbf{V}$ with row n replaced by $(1, c, \dots, c^{n-1})$. Then $\mathbf{V}_c$ has the same cofactors along row n as $\mathbf{V}$, so by Theorem 7.6.1 along row n ,

$$\det \mathbf{V}_c = \sum_{j=1}^n c^{j-1} C_{nj} = f(c). \quad (*)$$

In particular $\det \mathbf{V} = f(x_n)$. The coefficient of t^{n-1} in f is $C_{nn} = (-1)^{2n} \det \mathbf{V}(x_1, \dots, x_{n-1})$, since deleting row n and column n of $\mathbf{V}$ leaves the Vandermonde matrix of the first $n-1$ nodes.

Case 1: two of $x_1, \dots, x_{n-1}$ are equal, say $x_p = x_q$ with $p < q < n$. Then rows p and q of $\mathbf{V}$ are equal. By Theorem 7.4.5, $\det \mathbf{V} = \det \mathbf{V}^\top$, and $\mathbf{V}^\top$ has two equal columns, so $\det \mathbf{V} = 0$ because the determinant is alternating (Definition 7.2.2). The product also vanishes, since it contains the factor $x_q - x_p = 0$.

Case 2: $x_1, \dots, x_{n-1}$ are distinct. For $i < n$, the matrix $\mathbf{V}_{x_i}$ has rows i and n equal, so $f(x_i) = \det \mathbf{V}_{x_i} = 0$ by $(*)$ and the argument of Case 1. Hence $x_1, \dots, x_{n-1}$ are $n-1$ distinct roots of f , and by Corollary 6.1.4,

$$g(t) := (t - x_1)(t - x_2) \cdots (t - x_{n-1}) \text{ divides } f.$$

Write $f = hg$ with $h \in F[t]$. If $h \neq 0$, then $\deg h = \deg f - (n-1) \leq 0$ by Theorem 0.8.3, since $\deg f \leq n-1$; so in every case h is a constant $h \in F$. As g is monic of degree $n-1$, the coefficient of t^{n-1} in hg is h , so $h = C_{nn} = \det \mathbf{V}(x_1, \dots, x_{n-1})$. Evaluating at x_n (Theorem 0.8.5) and using the induction hypothesis,

$$\begin{aligned} \det \mathbf{V} &= f(x_n) = \det \mathbf{V}(x_1, \dots, x_{n-1}) \prod_{i=1}^{n-1} (x_n - x_i) \\ &= \prod_{1 \leq i < j \leq n-1} (x_j - x_i) \prod_{i=1}^{n-1} (x_n - x_i). \end{aligned}$$

The two products together run over all pairs $i < j \leq n$, the second one collecting exactly the pairs with $j = n$. This proves the formula for n , and the theorem follows by induction. ■

Example. A 3×3 Vandermonde Determinant

Compute $\det \mathbf{V}(1, 2, 4)$ from the formula and by cofactor expansion.

Solution. The formula gives $(2-1)(4-1)(4-2) = 1 \cdot 3 \cdot 2 = 6$. Directly, expanding

$$\mathbf{V}(1, 2, 4) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 4 & 16 \end{pmatrix}$$

along the first row gives $1 \cdot (32 - 16) - 1 \cdot (16 - 4) + 1 \cdot (4 - 2) = 16 - 12 + 2 = 6$.

The product vanishes exactly when a factor does, because a field has no zero divisors. So:

Corollary 7.8.2. Distinct Nodes Give an Invertible Vandermonde Matrix

$\mathbf{V}(x_1, \dots, x_n)$ is invertible if and only if $x_1, \dots, x_n$ are **distinct**.

Proof. In a field, a product is nonzero if and only if every factor is nonzero (Theorem 0.7.1 (f), applied repeatedly). So by Theorem 7.8.1, $\det \mathbf{V}(x_1, \dots, x_n) \neq 0$ if and only if $x_j \neq x_i$ for all $i < j$. By Theorem 7.5.2, this is equivalent to invertibility. ■

Now the promised two-line proof of Theorem 3.7.3. Given distinct $x_0, \dots, x_n$ and values $y_0, \dots, y_n$, a polynomial $p = c_0 + c_1x + \dots + c_nx^n$ satisfies $p(x_i) = y_i$ for all i if and only if $\mathbf{V}(x_0, \dots, x_n) \mathbf{c} = \mathbf{y}$, since row i of this system reads $c_0 + c_1x_i + \dots + c_nx_i^n = y_i$. By Corollary 7.8.2 the matrix is invertible, so the system has exactly one solution $\mathbf{c} = \mathbf{V}(x_0, \dots, x_n)^{-1} \mathbf{y}$ (Theorem 3.4.4), and the interpolating polynomial exists and is unique.

Remark. This is not a proof from nothing. The root bound of Chapter 6 (Corollary 6.1.4) did the real work inside Theorem 7.8.1, exactly as the root bound did in Chapter 3 (Lemma 3.7.2). What the determinant adds is an explicit number: it says how close to singular the interpolation problem is, since nodes that nearly coincide make the product small.

Warning. Watch the shape and the order. The formula is for rows $(1, x_i, \dots, x_i^{n-1})$ starting at the power 0, and for the factors $x_j - x_i$ with $i < j$, “later node minus earlier node”. Dropping the column of ones changes the answer: $\det \begin{pmatrix} x_1 & x_1^2 \\ x_2 & x_2^2 \end{pmatrix} = x_1x_2(x_2 - x_1)$, not $x_2 - x_1$. Writing the factors as $x_i - x_j$ instead flips the sign of each of the $\binom{n}{2}$ factors, so it multiplies the answer by $(-1)^{\binom{n}{2}}$: already for $n = 2$ and $n = 3$ this is -1 , a wrong sign.

Quick check. Without expanding, find $\det \mathbf{V}(1, 2, 3, 5)$.

Solution. The pairs $i < j$ give $(2-1)(3-1)(5-1)(3-2)(5-2)(5-3) = 1 \cdot 2 \cdot 4 \cdot 1 \cdot 3 \cdot 2 = 48$.

7.8.2 Tridiagonal determinants

Many matrices in applications have nonzero entries only on the diagonal and next to it: discretized differential equations, chains of springs, the recurrences of orthogonal polynomials. Such a matrix is **tridiagonal**. Its determinant is not a closed formula in general, but it obeys

a recurrence in the size, because the last row has only two nonzero entries.

Let $\mathbf{T}_n \in M_n(F)$ have diagonal entries $a_1, \dots, a_n$, entries $b_1, \dots, b_{n-1}$ just above the diagonal ($(i, i+1)$ -entry b_i), entries $c_1, \dots, c_{n-1}$ just below it ($(i+1, i)$ -entry c_i), and zeros elsewhere:

$$\mathbf{T}_n = \begin{pmatrix} a_1 & b_1 & & & \\ c_1 & a_2 & b_2 & & \\ & c_2 & \ddots & \ddots & \\ & & \ddots & a_{n-1} & b_{n-1} \\ & & & c_{n-1} & a_n \end{pmatrix}.$$

For $1 \leq k \leq n$, the top-left $k \times k$ block of $\mathbf{T}_n$ is the tridiagonal matrix $\mathbf{T}_k$ built from $a_1, \dots, a_k$, $b_1, \dots, b_{k-1}$, $c_1, \dots, c_{k-1}$. Write $D_k = \det \mathbf{T}_k$, and set $D_0 = 1$.

Theorem 7.8.3. Tridiagonal Recurrence

With this notation, $D_1 = a_1$, and for $2 \leq k \leq n$,

$$D_k = a_k D_{k-1} - b_{k-1} c_{k-1} D_{k-2}.$$

Proof. $D_1 = \det(a_1) = a_1$. For $k = 2$, $D_2 = a_1 a_2 - b_1 c_1 = a_2 D_1 - b_1 c_1 D_0$. Let $k \geq 3$. Row k of $\mathbf{T}_k$ has only two possibly nonzero entries, c_{k-1} in column $k-1$ and a_k in column k . By Theorem 7.6.1 along row k ,

$$D_k = c_{k-1}(-1)^{2k-1} M_{k,k-1} + a_k(-1)^{2k} M_{kk} = a_k M_{kk} - c_{k-1} M_{k,k-1},$$

where $M_{kk} = \det \mathbf{T}_{k-1} = D_{k-1}$. The minor $M_{k,k-1}$ is the determinant of the $(k-1) \times (k-1)$ matrix obtained by deleting row k and column $k-1$. Its last column consists of rows $1, \dots, k-1$ of column k of $\mathbf{T}_k$, which is b_{k-1} in the last position and 0 elsewhere. Expanding along this last column (Theorem 7.6.1), with sign $(-1)^{(k-1)+(k-1)} = 1$, gives $M_{k,k-1} = b_{k-1} \det \mathbf{T}_{k-2} = b_{k-1} D_{k-2}$, because deleting the last row and column leaves rows and columns $1, \dots, k-2$ of $\mathbf{T}_k$. Substituting proves the recurrence. ■

Example. Fibonacci Numbers as Determinants

Let $\mathbf{T}_n$ have 1 on the diagonal, 1 above it, and -1 below it. Find $D_n = \det \mathbf{T}_n$.

Solution. Here $a_k = 1$ and $b_{k-1} c_{k-1} = 1 \cdot (-1) = -1$, so Theorem 7.8.3 gives $D_k = D_{k-1} + D_{k-2}$ for $k \geq 2$, with $D_0 = 1$ and $D_1 = 1$. Hence

$$D_1, D_2, D_3, \dots = 1, 2, 3, 5, 8, 13, 21, 34, \dots,$$

the Fibonacci numbers: $D_n = F_{n+1}$, where $F_1 = F_2 = 1$ and $F_k = F_{k-1} + F_{k-2}$. For instance

$$D_3 = \det \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 1 \\ 0 & -1 & 1 \end{pmatrix} = 1 \cdot (1 + 1) - 1 \cdot (-1 - 0) + 0 = 3.$$

Only the products $b_{k-1}c_{k-1}$ enter the recurrence, never b_{k-1} and c_{k-1} separately. For example, for a real number t , putting $2t$ on the diagonal and 1 on both neighboring lines gives $D_k = 2t D_{k-1} - D_{k-2}$, so $D_1 = 2t$, $D_2 = 4t^2 - 1$, $D_3 = 8t^3 - 4t$. These are the Chebyshev polynomials of the second kind.

7.8.3 Block triangular matrices

A matrix can be cut into blocks by a horizontal and a vertical line. When the block in the lower left is zero, the determinant should see only the two square blocks on the diagonal, just as a triangular determinant sees only diagonal entries. This result is used constantly, for instance in Chapter 8 for Schur complements.

Theorem 7.8.4. Block Triangular Determinant

Let $\mathbf{A} \in M_k(F)$, $\mathbf{D} \in M_l(F)$ and $\mathbf{B} \in M_{k \times l}(F)$, with $k, l \geq 1$. Then

$$\det \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{O} & \mathbf{D} \end{pmatrix} = \det \mathbf{A} \cdot \det \mathbf{D},$$

where $\mathbf{O}$ is the $l \times k$ zero matrix. The same holds for $\begin{pmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}$ with $\mathbf{C} \in M_{l \times k}(F)$.

Idea. In the Leibniz formula for the whole matrix, a term picks one entry from each column. The first k columns have zeros below row k , so a nonzero term must pick rows $1, \dots, k$ for the first k columns, and is then forced to pick rows $k+1, \dots, n$ for the last l columns. Such a permutation is a permutation of each block, its inversions are the inversions inside the blocks, and the Leibniz sum factors into two Leibniz sums.

Proof. Let $n = k + l$ and $\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{O} & \mathbf{D} \end{pmatrix} = (m_{ij})$. By Definition 7.4.1, $\det \mathbf{M} = \sum_{\sigma \in S_n} \text{sgn}(\sigma) \prod_{j=1}^n m_{\sigma(j)j}$. Since $m_{ij} = 0$ for $i > k \geq j$, the term of σ is 0 unless $\sigma(j) \leq k$ for every $j \leq k$. For such σ , the injective map σ sends $\{1, \dots, k\}$ into itself, hence onto itself (a finite set), and so it sends $\{k+1, \dots, n\}$ onto itself. Therefore σ determines $\alpha \in S_k$ and $\beta \in S_l$ by

$$\alpha(j) = \sigma(j) \quad (j \leq k), \quad \beta(q) = \sigma(k+q) - k \quad (q \leq l),$$

and every pair (α, β) arises from exactly one such σ .

We compare signs through inversions (Definition 7.3.1). A pair $i < j$ with $i, j \leq k$ is an inversion of σ exactly when it is one of α ; a pair $k+p < k+q$ is an inversion of σ exactly when $p < q$ is one of β ; and a pair $i \leq k < j$ is never an inversion, since $\sigma(i) \leq k < \sigma(j)$. By Definition 7.3.2, $\text{sgn}(\sigma) = \text{sgn}(\alpha) \text{sgn}(\beta)$. Also $m_{\sigma(j)j} = a_{\alpha(j)j}$ for $j \leq k$

and $m_{\sigma(k+q),k+q} = d_{\beta(q)q}$. Hence

$$\det \mathbf{M} = \sum_{\alpha \in S_k} \sum_{\beta \in S_l} \operatorname{sgn}(\alpha) \operatorname{sgn}(\beta) \prod_{j=1}^k a_{\alpha(j)j} \prod_{q=1}^l d_{\beta(q)q} = \det \mathbf{A} \cdot \det \mathbf{D},$$

using Definition 7.4.1 twice. For the lower block triangular case, the transpose of $\begin{pmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}$ is $\begin{pmatrix} \mathbf{A}^\top & \mathbf{C}^\top \\ \mathbf{O} & \mathbf{D}^\top \end{pmatrix}$, so by Theorem 7.4.5 and the first case its determinant is $\det \mathbf{A}^\top \det \mathbf{D}^\top = \det \mathbf{A} \det \mathbf{D}$. ■

The proof uses only sums and products of entries, so, like the results flagged in §4–§6, it holds for matrices over any commutative ring; §10 uses this over $F[x]$. By induction on the number of blocks, a matrix with square blocks $\mathbf{A}_1, \dots, \mathbf{A}_r$ on the diagonal and zeros below them has determinant $\det \mathbf{A}_1 \cdots \det \mathbf{A}_r$; Theorem 7.4.6 is the case where every block is 1×1 .

For example,

$$\det \begin{pmatrix} 2 & 1 & 5 & -3 \\ 1 & 1 & 0 & 7 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 4 & 2 \end{pmatrix} = \det \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \det \begin{pmatrix} 3 & 1 \\ 4 & 2 \end{pmatrix} = 1 \cdot 2 = 2,$$

and the entries 5, −3, 0, 7 of the upper right block play no role.

Warning. **There is no “block $ad - bc$ ” rule.** For a general block matrix, $\det \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \neq \det \mathbf{A} \det \mathbf{D} - \det \mathbf{B} \det \mathbf{C}$. Take $\mathbf{A} = \mathbf{D} = \mathbf{O}$ and $\mathbf{B} = \mathbf{C} = \mathbf{I}_2$: the 4×4 matrix $\begin{pmatrix} \mathbf{O} & \mathbf{I}_2 \\ \mathbf{I}_2 & \mathbf{O} \end{pmatrix}$ is the permutation matrix of $(1\ 3)(2\ 4)$, a product of two transpositions, so its determinant is +1; the rule would give $0 \cdot 0 - 1 \cdot 1 = -1$. The zero block in the corner is what the theorem needs. The correct formulas for general blocks use Schur complements, in Chapter 8.

Quick check. Is $\det \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{O} & \mathbf{D} \end{pmatrix} = \det \mathbf{A} \det \mathbf{D}$ meaningful when $\mathbf{A}$ is 2×3 and $\mathbf{D}$ is 3×2 (so the whole matrix is 5×5)?

Solution. No. The theorem requires $\mathbf{A}$ and $\mathbf{D}$ square; here $\det \mathbf{A}$ and $\det \mathbf{D}$ are undefined. In fact the whole determinant is 0 whatever the entries: the first 3 columns have their nonzero entries only in the first 2 rows, so they are linearly dependent in a 2-dimensional space, and a determinant with dependent columns is 0 (Theorem 7.2.1).

7.8.4 Rank-one updates

Changing a matrix by a rank-one matrix $\mathbf{u}\mathbf{v}^\top$ is the smallest possible change in the sense of rank, and it happens often: updating one observation in statistics, modifying one row of a system, or adding the same number to every entry. The determinant of the updated identity turns out to be a single dot product.

Theorem 7.8.5. Matrix Determinant Lemma

Let $\mathbf{u}, \mathbf{v} \in F^n$. Then

$$\det(\mathbf{I}_n + \mathbf{u}\mathbf{v}^\top) = 1 + \mathbf{v}^\top \mathbf{u}.$$

Idea. Border $\mathbf{I}_n + \mathbf{u}\mathbf{v}^\top$ by an extra column $\mathbf{u}$ and an extra row $(\mathbf{0}^\top, 1)$, giving a block triangular matrix $\mathbf{M}$ with the same determinant. Then clear the update by elimination with blocks. Rows: add $\mathbf{v}^\top$ times the top n rows to the new bottom row (multiply by $\mathbf{L}$ on the left). Columns: subtract the last column times $\mathbf{v}^\top$ from the first n columns (multiply by $\mathbf{L}'$ on the right). The factors $\mathbf{L}, \mathbf{L}'$ have determinant 1, and the result is block triangular with $\mathbf{I}_n$ and the 1×1 block $1 + \mathbf{v}^\top \mathbf{u}$ on the diagonal.

Proof. Consider the $(n+1) \times (n+1)$ block matrices

$$\mathbf{L} = \begin{pmatrix} \mathbf{I}_n & \mathbf{0} \\ \mathbf{v}^\top & 1 \end{pmatrix}, \quad \mathbf{M} = \begin{pmatrix} \mathbf{I}_n + \mathbf{u}\mathbf{v}^\top & \mathbf{u} \\ \mathbf{0}^\top & 1 \end{pmatrix}, \quad \mathbf{L}' = \begin{pmatrix} \mathbf{I}_n & \mathbf{0} \\ -\mathbf{v}^\top & 1 \end{pmatrix}.$$

Splitting the sum $\sum_j$ in Theorem 0.9.2 (1) into the indices $j \leq n$ and $j = n+1$ shows that matrices partitioned in this way multiply block by block, like 2×2 matrices whose entries are blocks. Hence

$$\mathbf{L}\mathbf{M} = \begin{pmatrix} \mathbf{I}_n + \mathbf{u}\mathbf{v}^\top & \mathbf{u} \\ \mathbf{v}^\top + (\mathbf{v}^\top \mathbf{u})\mathbf{v}^\top & \mathbf{v}^\top \mathbf{u} + 1 \end{pmatrix},$$

where we used $\mathbf{v}^\top(\mathbf{u}\mathbf{v}^\top) = (\mathbf{v}^\top \mathbf{u})\mathbf{v}^\top$ by associativity, $\mathbf{v}^\top \mathbf{u}$ being a 1×1 matrix, that is, a scalar. Multiplying on the right by $\mathbf{L}'$,

$$\mathbf{L}\mathbf{M}\mathbf{L}' = \begin{pmatrix} \mathbf{I}_n + \mathbf{u}\mathbf{v}^\top - \mathbf{u}\mathbf{v}^\top & \mathbf{u} \\ \mathbf{v}^\top + (\mathbf{v}^\top \mathbf{u})\mathbf{v}^\top - (1 + \mathbf{v}^\top \mathbf{u})\mathbf{v}^\top & 1 + \mathbf{v}^\top \mathbf{u} \end{pmatrix} = \begin{pmatrix} \mathbf{I}_n & \mathbf{u} \\ \mathbf{0}^\top & 1 + \mathbf{v}^\top \mathbf{u} \end{pmatrix}.$$

By Theorem 7.8.4, $\det \mathbf{L} = \det \mathbf{I}_n \cdot 1 = 1$, $\det \mathbf{L}' = 1$, $\det \mathbf{M} = \det(\mathbf{I}_n + \mathbf{u}\mathbf{v}^\top) \cdot 1$, and $\det(\mathbf{L}\mathbf{M}\mathbf{L}') = \det \mathbf{I}_n (1 + \mathbf{v}^\top \mathbf{u}) = 1 + \mathbf{v}^\top \mathbf{u}$. By Theorem 7.5.4, $\det(\mathbf{L}\mathbf{M}\mathbf{L}') = \det \mathbf{L} \det \mathbf{M} \det \mathbf{L}' = \det(\mathbf{I}_n + \mathbf{u}\mathbf{v}^\top)$. This proves the lemma. ■

Example. a on the Diagonal, 1 Elsewhere

Let $a \in \mathbb{R}$ and $n \geq 1$, and let $\mathbf{A}_n \in M_n(\mathbb{R})$ have every diagonal entry a and every other entry 1. Find $\det \mathbf{A}_n$.

Solution. Let $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^n$. Every entry of $\mathbf{1}\mathbf{1}^\top$ is 1, so $\mathbf{A}_n = (a-1)\mathbf{I}_n + \mathbf{1}\mathbf{1}^\top$.

Case $a \neq 1$. Then $\mathbf{A}_n = (a-1)(\mathbf{I}_n + \mathbf{u}\mathbf{1}^\top)$ with $\mathbf{u} = \frac{1}{a-1}\mathbf{1}$. Scaling an $n \times n$ matrix by c multiplies its determinant by c^n (multilinearity in each of the n columns), so by [Theorem 7.8.5](#)

$$\det \mathbf{A}_n = (a-1)^n \left(1 + \frac{n}{a-1}\right) = (a-1)^{n-1}(a-1+n),$$

since $\mathbf{1}^\top \mathbf{u} = \frac{n}{a-1}$.

Case $a = 1$. For $n \geq 2$, all columns are equal, so $\det \mathbf{A}_n = 0$, which the formula also gives. For $n = 1$, $\det \mathbf{A}_1 = 1$, and the formula gives $(a-1)^0 \cdot 1 = 1$.

So $\det \mathbf{A}_n = (a-1)^{n-1}(a+n-1)$ for every a . For $n = 3$ this is $(a-1)^2(a+2)$; with $a = 3$ it gives $4 \cdot 5 = 20$.

Warning. The lemma is about $\mathbf{I}_n + \mathbf{u}\mathbf{v}^\top$, not $\mathbf{u}\mathbf{v}^\top$ or $\mathbf{u}^\top \mathbf{v}$. The rank-one matrix $\mathbf{u}\mathbf{v}^\top$ is $n \times n$ and, for $n \geq 2$, has determinant 0; the scalar $\mathbf{v}^\top \mathbf{u}$ is its trace. Do not confuse $\det(\mathbf{I}_n + \mathbf{u}\mathbf{v}^\top) = 1 + \mathbf{v}^\top \mathbf{u}$ with $\det(\mathbf{I}_n) + \det(\mathbf{u}\mathbf{v}^\top) = 1$ (for $n \geq 2$), which is false whenever $\mathbf{v}^\top \mathbf{u} \neq 0$: for $\mathbf{u} = \mathbf{v} = \mathbf{e}_1$, $\mathbf{I}_n + \mathbf{e}_1\mathbf{e}_1^\top = \text{diag}(2, 1, \dots, 1)$ has determinant 2.

7.8.5 Exercises

7.8.5.1 A. Check your understanding

Exercise. A1

- State the Vandermonde determinant formula, including the order of the factors.
- True or false: $\mathbf{V}(x_1, \dots, x_n)$ is invertible whenever $x_1, \dots, x_n$ are nonzero. Justify your answer.
- Write down the recurrence for the determinant of a tridiagonal matrix, and say which row the proof expands along.
- True or false: $\det \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \det \mathbf{A} \det \mathbf{D}$ whenever $\mathbf{A}$ and $\mathbf{D}$ are square and $\mathbf{B} = \mathbf{O}$. Justify your answer.
- True or false: $\det(\mathbf{I}_n + \mathbf{u}\mathbf{v}^\top) = \det(\mathbf{I}_n + \mathbf{v}\mathbf{u}^\top)$ for all $\mathbf{u}, \mathbf{v} \in F^n$. Justify your answer.

Solution.

- $\det \mathbf{V}(x_1, \dots, x_n) = \prod_{1 \leq i < j \leq n} (x_j - x_i)$, where $\mathbf{V}$ has (i, j) -entry x_i^{j-1} ([Theorem 7.8.1](#)).
- False. $\mathbf{V}(1, 1) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ has nonzero nodes and determinant 0. The condition is that the nodes are **distinct** ([Corollary 7.8.2](#)); 0 is allowed as a node.
- $D_k = a_k D_{k-1} - b_{k-1} c_{k-1} D_{k-2}$, with $D_0 = 1$, $D_1 = a_1$ ([Theorem 7.8.3](#)). The proof expands along the last row, which has only two nonzero entries.
- True. With $\mathbf{B} = \mathbf{O}$ the matrix is lower block triangular, and [Theorem 7.8.4](#) applies.

(e) True. Both equal $1 + \mathbf{v}^\top \mathbf{u} = 1 + \mathbf{u}^\top \mathbf{v}$ by [Theorem 7.8.5](#), since $\mathbf{v}^\top \mathbf{u} = \sum_i v_i u_i = \mathbf{u}^\top \mathbf{v}$.

7.8.5.2 B. Practice

Exercise. B1: A 4×4 Vandermonde determinant

Compute $\det \begin{pmatrix} 1 & -1 & 1 & -1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \end{pmatrix}$. Hence decide whether there is exactly one polynomial $p \in \mathbb{R}[x]_{\leq 3}$ with $p(-1) = 3$, $p(0) = 1$, $p(1) = 0$, $p(2) = 5$.

Solution. Row i is $(1, x_i, x_i^2, x_i^3)$ with nodes $-1, 0, 1, 2$, so this is $\mathbf{V}(-1, 0, 1, 2)$. By [Theorem 7.8.1](#) its determinant is

$$(0 - (-1))(1 - (-1))(2 - (-1))(1 - 0)(2 - 0)(2 - 1) = 1 \cdot 2 \cdot 3 \cdot 1 \cdot 2 \cdot 1 = 12.$$

(Expand along row 2, which is $(1, 0, 0, 0)$: the determinant is $(-1)^{2+1} \det \begin{pmatrix} -1 & 1 & -1 \\ 1 & 1 & 1 \\ 2 & 4 & 8 \end{pmatrix}$,

that is $-(-1(8 - 4) - 1(8 - 2) - 1(4 - 2)) = -(-4 - 6 - 2) = 12$.) Hence the coefficient matrix of the interpolation system is invertible, and there is exactly one $p \in \mathbb{R}[x]_{\leq 3}$ with the given values.

Exercise. B2: The second-difference matrix

Let $\mathbf{K}_n \in M_n(\mathbb{R})$ have 2 on the diagonal, -1 directly above and below it, and 0 elsewhere. Prove that $\det \mathbf{K}_n = n + 1$. Hence show that $\mathbf{K}_n$ is invertible for every n .

Solution. Here $a_k = 2$ and $b_{k-1}c_{k-1} = (-1)(-1) = 1$. By [Theorem 7.8.3](#), $D_k = 2D_{k-1} - D_{k-2}$ for $k \geq 2$, with $D_0 = 1$, $D_1 = 2$. We show $D_k = k + 1$ by strong induction on k ([Theorem 0.2.10](#)). It holds for $k = 0, 1$. If it holds for $k-1$ and $k-2$, then $D_k = 2k - (k-1) = k + 1$. Hence $\det \mathbf{K}_n = D_n = n + 1 \neq 0$, and $\mathbf{K}_n$ is invertible by [Theorem 7.5.2](#).

Exercise. B3: A rank-one update

Let $\mathbf{u} = (1, 2, 3)$ and $\mathbf{v} = (1, -1, 2)$ in $\mathbb{Q}^3$.

- Use [Theorem 7.8.5](#) to find $\det(\mathbf{I}_3 + \mathbf{u}\mathbf{v}^\top)$.
- Write out $\mathbf{I}_3 + \mathbf{u}\mathbf{v}^\top$ and check your answer by elimination.
- Find all $t \in \mathbb{Q}$ for which $\mathbf{I}_3 + t\mathbf{u}\mathbf{v}^\top$ is **not** invertible.

Solution.

- $\mathbf{v}^\top \mathbf{u} = 1 - 2 + 6 = 5$, so $\det(\mathbf{I}_3 + \mathbf{u}\mathbf{v}^\top) = 6$.

- (b) $\mathbf{uv}^\top = \begin{pmatrix} 1 & -1 & 2 \\ 2 & -2 & 4 \\ 3 & -3 & 6 \end{pmatrix}$, so $\mathbf{I}_3 + \mathbf{uv}^\top = \begin{pmatrix} 2 & -1 & 2 \\ 2 & -1 & 4 \\ 3 & -3 & 7 \end{pmatrix}$. Subtract row 1 from row 2, which does not change the determinant (Theorem 7.5.1), to get row 2 equal to $(0, 0, 2)$. Expanding along that row,

$$\det = 2 \cdot (-1)^{2+3} \det \begin{pmatrix} 2 & -1 \\ 3 & -3 \end{pmatrix} = -2(-6 + 3) = 6.$$

- (c) $\mathbf{I}_3 + t\mathbf{uv}^\top = \mathbf{I}_3 + (t\mathbf{u})\mathbf{v}^\top$, whose determinant is $1 + t\mathbf{v}^\top\mathbf{u} = 1 + 5t$. It is zero exactly for $t = -\frac{1}{5}$, and by Theorem 7.5.2 that is the only t for which the matrix is not invertible.

7.8.5.3 C. Going deeper

Exercise. C1: The Cauchy determinant

Let $x_1, \dots, x_n, y_1, \dots, y_n \in F$ with $x_i + y_j \neq 0$ for all i, j , and let $\mathbf{Q}_n \in M_n(F)$ have (i, j) -entry $\frac{1}{x_i + y_j}$. The goal is

$$\det \mathbf{Q}_n = \frac{\prod_{1 \leq i < j \leq n} (x_j - x_i)(y_j - y_i)}{\prod_{i=1}^n \prod_{j=1}^n (x_i + y_j)}.$$

- (a) For $i < n$, subtract row n from row i . Show that the new (i, j) -entry is $\frac{x_n - x_i}{(x_i + y_j)(x_n + y_j)}$, and deduce that $\det \mathbf{Q}_n = \frac{\prod_{i < n} (x_n - x_i)}{\prod_j (x_n + y_j)} \det \mathbf{Q}'$, where $\mathbf{Q}'$ agrees with $\mathbf{Q}_n$ in rows $1, \dots, n-1$ and has last row $(1, \dots, 1)$.
- (b) In $\mathbf{Q}'$, for $j < n$, subtract column n from column j . Deduce that $\det \mathbf{Q}' = \frac{\prod_{j < n} (y_n - y_j)}{\prod_{i < n} (x_i + y_n)} \det \mathbf{Q}_{n-1}$, where $\mathbf{Q}_{n-1}$ is built from $x_1, \dots, x_{n-1}$ and $y_1, \dots, y_{n-1}$.
- (c) Prove the formula by induction on n .
- (d) The **Hilbert matrix** $\mathbf{H}_n$ has (i, j) -entry $\frac{1}{i+j-1}$. Show that it is a Cauchy matrix, and use (c) to compute $\det \mathbf{H}_3$.

Hint: in (a) and (b), factor scalars out of whole rows and columns using multilinearity.

Solution.

- (a) For $i < n$, $\frac{1}{x_i + y_j} - \frac{1}{x_n + y_j} = \frac{(x_n + y_j) - (x_i + y_j)}{(x_i + y_j)(x_n + y_j)} = \frac{x_n - x_i}{(x_i + y_j)(x_n + y_j)}$. These row operations do not change the determinant (Theorem 7.5.1). In the new matrix, row $i < n$ has the factor $x_n - x_i$ in every entry, and column j has the factor $\frac{1}{x_n + y_j}$ in every entry (in row n , $\frac{1}{x_n + y_j} = \frac{1}{x_n + y_j} \cdot 1$). Pulling these scalars out of each row and column by multilinearity (Theorem 7.5.1, in its row and column forms) leaves the matrix $\mathbf{Q}'$ with entries $\frac{1}{x_i + y_j}$ in rows $i < n$ and 1 in row n . This gives the stated formula.
- (b) For $j < n$, subtracting column n from column j leaves the determinant unchanged, makes the last row $(0, \dots, 0, 1)$, and for $i < n$ gives $\frac{1}{x_i + y_j} - \frac{1}{x_i + y_n} = \frac{y_n - y_j}{(x_i + y_j)(x_i + y_n)}$. Pull $y_n - y_j$ out of column j ($j < n$) and $\frac{1}{x_i + y_n}$ out of row i ($i < n$); the last row $(0, \dots, 0, 1)$ is unaffected by the column factors because its first $n-1$ entries are 0,

and row factors do not touch row n . What remains has $\frac{1}{x_i+y_j}$ in positions $i, j < n$, entries $(x_i + y_n) \cdot \frac{1}{x_i+y_n} = 1$ in column n above the corner, and last row $(0, \dots, 0, 1)$. Expanding along the last row (Theorem 7.6.1) gives $\det \mathbf{Q}_{n-1}$, with sign $(-1)^{2n} = 1$.

- (c) For $n = 1$, $\det \mathbf{Q}_1 = \frac{1}{x_1+y_1}$, which is the formula with empty products upstairs. For $n \geq 2$, combining (a) and (b),

$$\det \mathbf{Q}_n = \frac{\prod_{i < n} (x_n - x_i)(y_n - y_i)}{\prod_{j \leq n} (x_n + y_j) \prod_{i < n} (x_i + y_n)} \det \mathbf{Q}_{n-1}.$$

The numerator supplies exactly the factors of $\prod_{i < j} (x_j - x_i)(y_j - y_i)$ with $j = n$, and the denominator supplies exactly the factors $x_i + y_j$ with $\max(i, j) = n$, each once. By the induction hypothesis, $\det \mathbf{Q}_{n-1}$ supplies all the remaining factors. This proves the formula.

- (d) With $x_i = i$ and $y_j = j - 1$, $x_i + y_j = i + j - 1 \geq 1$, so $\mathbf{H}_n$ is a Cauchy matrix over $\mathbb{Q}$. For $n = 3$, the numerator is $((2-1)(3-1)(3-2))^2 = 4$, since $y_j - y_i = j - i = x_j - x_i$. The denominator is the product of all $i + j - 1$: the rows give $1 \cdot 2 \cdot 3 = 6$, $2 \cdot 3 \cdot 4 = 24$, $3 \cdot 4 \cdot 5 = 60$, in total 8640. Hence $\det \mathbf{H}_3 = \frac{4}{8640} = \frac{1}{2160}$. (In general $\det \mathbf{H}_n = c_n^4 / c_{2n}$ with $c_n = \prod_{k=1}^{n-1} k!$; we state this without proof. These determinants shrink extremely fast, which is why Hilbert matrices are a standard test of numerical methods.)

Exercise. C2: A circulant determinant over $\mathbb{C}$

Let $a, b, c \in \mathbb{C}$, let $\omega = e^{2\pi i/3}$, and let

$$\mathbf{Z} = \begin{pmatrix} a & b & c \\ c & a & b \\ b & c & a \end{pmatrix}.$$

- (a) For $k = 0, 1, 2$, let $\mathbf{w}_k = (1, \omega^k, \omega^{2k})$ and $\lambda_k = a + b\omega^k + c\omega^{2k}$. Show that $\mathbf{Z}\mathbf{w}_k = \lambda_k \mathbf{w}_k$.
 (b) Let $\mathbf{W}$ be the matrix with columns $\mathbf{w}_0, \mathbf{w}_1, \mathbf{w}_2$. Show that $\mathbf{Z}\mathbf{W} = \mathbf{W} \operatorname{diag}(\lambda_0, \lambda_1, \lambda_2)$ and that $\det \mathbf{W} \neq 0$. Deduce $\det \mathbf{Z} = \lambda_0 \lambda_1 \lambda_2$.
 (c) Using $\omega^3 = 1$ and $1 + \omega + \omega^2 = 0$, deduce $\det \mathbf{Z} = a^3 + b^3 + c^3 - 3abc$, and check this against the Leibniz formula.

Circulant matrices of every size are treated in Chapter 12.

Solution.

- (a) By Example (Roots of Unity), $\omega^3 = 1$, so $\omega^{-k} = \omega^{2k}$ and $\omega^{-2k} = \omega^k$. Row 1 of $\mathbf{Z}\mathbf{w}_k$ is $a + b\omega^k + c\omega^{2k} = \lambda_k$. Row 2 is $c + a\omega^k + b\omega^{2k} = \omega^k(c\omega^{-k} + a + b\omega^k) = \omega^k(a + b\omega^k + c\omega^{2k}) = \omega^k \lambda_k$. Row 3 is $b + c\omega^k + a\omega^{2k} = \omega^{2k}(b\omega^{-2k} + c\omega^{-k} + a) = \omega^{2k} \lambda_k$. Hence $\mathbf{Z}\mathbf{w}_k = \lambda_k \mathbf{w}_k$.
 (b) By Theorem 0.9.2, column k of $\mathbf{Z}\mathbf{W}$ is $\mathbf{Z}\mathbf{w}_k = \lambda_k \mathbf{w}_k$, column k of $\mathbf{W} \operatorname{diag}(\lambda_0, \lambda_1, \lambda_2)$. The (i, k) -entry of $\mathbf{W}$ is $(\omega^k)^{i-1}$, so $\mathbf{W} = \mathbf{V}(1, \omega, \omega^2)^\top$. The nodes $1, \omega, \omega^2$ are distinct (they are $e^{2\pi i k/3}$ for $k = 0, 1, 2$), so $\det \mathbf{W} = \det \mathbf{V}(1, \omega, \omega^2) \neq 0$ by Theorem 7.4.5 and Corollary 7.8.2. By Theorem 7.5.4 and Theorem 7.4.6, $\det \mathbf{Z} \det \mathbf{W} = \det \mathbf{W} \lambda_0 \lambda_1 \lambda_2$, and canceling $\det \mathbf{W} \neq 0$ gives $\det \mathbf{Z} = \lambda_0 \lambda_1 \lambda_2$.
 (c) Multiply out $\lambda_1 \lambda_2 = (a + b\omega + c\omega^2)(a + b\omega^2 + c\omega)$. The nine terms are $a^2, ab\omega^2 + ab\omega,$

$ac\omega + ac\omega^2, b^2\omega^3, bc\omega^2, bc\omega^4 = bc\omega, c^2\omega^3$. Using $\omega^3 = 1$ and $\omega + \omega^2 = -1$,

$$\lambda_1\lambda_2 = a^2 + b^2 + c^2 - ab - ac - bc.$$

Then $\det \mathbf{Z} = (a + b + c)(a^2 + b^2 + c^2 - ab - ac - bc) = a^3 + b^3 + c^3 - 3abc$, expanding and canceling. The Leibniz formula for $\mathbf{Z}$ gives the three even terms $a \cdot a \cdot a$, $c \cdot c \cdot c$ (from $\sigma = (1\ 2\ 3)$), and $b \cdot b \cdot b$, and the three odd terms, each equal to abc with sign $-$: the same answer.

Exercise. C3: Rank-one updates of an invertible matrix

Let $\mathbf{A} \in M_n(F)$ be invertible and $\mathbf{u}, \mathbf{v} \in F^n$.

- Prove that $\det(\mathbf{A} + \mathbf{u}\mathbf{v}^\top) = \det \mathbf{A} (1 + \mathbf{v}^\top \mathbf{A}^{-1} \mathbf{u})$.
- Deduce that $\mathbf{A} + \mathbf{u}\mathbf{v}^\top$ is invertible if and only if $\mathbf{v}^\top \mathbf{A}^{-1} \mathbf{u} \neq -1$.
- Give an example over $\mathbb{Q}$ with $n = 2$ where $\mathbf{A}$ is invertible, $\mathbf{u}\mathbf{v}^\top \neq \mathbf{O}$, and $\mathbf{A} + \mathbf{u}\mathbf{v}^\top$ is **not** invertible.

Solution.

- Since $\mathbf{A}$ is invertible, $\mathbf{A} + \mathbf{u}\mathbf{v}^\top = \mathbf{A}(\mathbf{I}_n + \mathbf{A}^{-1}\mathbf{u}\mathbf{v}^\top) = \mathbf{A}(\mathbf{I}_n + (\mathbf{A}^{-1}\mathbf{u})\mathbf{v}^\top)$ by Theorem 0.9.4. By Theorem 7.5.4 and Theorem 7.8.5 applied to the vectors $\mathbf{A}^{-1}\mathbf{u}$ and $\mathbf{v}$, $\det(\mathbf{A} + \mathbf{u}\mathbf{v}^\top) = \det \mathbf{A} (1 + \mathbf{v}^\top \mathbf{A}^{-1} \mathbf{u})$.
- $\det \mathbf{A} \neq 0$ by Theorem 7.5.2, so $\det(\mathbf{A} + \mathbf{u}\mathbf{v}^\top) \neq 0$ exactly when $1 + \mathbf{v}^\top \mathbf{A}^{-1} \mathbf{u} \neq 0$. Apply Theorem 7.5.2 again.
- Take $\mathbf{A} = \mathbf{I}_2$, $\mathbf{u} = (-1, 0)$, $\mathbf{v} = (1, 0)$. Then $\mathbf{v}^\top \mathbf{u} = -1$, and indeed $\mathbf{A} + \mathbf{u}\mathbf{v}^\top = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ is not invertible.