

21.9 Loewner Matrices

Chapter 17 §11 left operator monotonicity as an infinite list of conditions: one for every matrix size, each quantified over all pairs of Hermitian matrices of that size. Nothing in that list can be checked. This section replaces the condition at size n by a condition on **numbers**: a single symmetric $n \times n$ matrix built out of f alone, at n chosen points, must be positive semidefinite. Two ingredients do the work — a formula for the derivative of $s \mapsto f(\mathbf{A} + s\mathbf{H})$, and the Schur product theorem of Chapter 13 §08. They are the elementary half of Loewner's theorem, and they are proved here in full.

Throughout, $F = \mathbb{R}$ or $F = \mathbb{C}$, every matrix whose spectrum is named is Hermitian, and I denotes an **interval** of real numbers with more than one point. The letter I is plain italic and is not the identity matrix $\mathbf{I}$. For $f: I \rightarrow \mathbb{R}$ and Hermitian $\mathbf{A}$ with $\text{spec}(\mathbf{A}) \subseteq I$, the matrix $f(\mathbf{A})$ is the one built by the functional calculus of Chapter 12 §08 (Definition 12.8.2). We write $f \in C^1(I)$ to mean that f is differentiable on I , with one-sided derivatives at any endpoint belonging to I , and that f' is continuous on I .

21.9.1 Slopes between two points

Operator monotonicity compares $f(\mathbf{A})$ with $f(\mathbf{B})$, and the difference of two values of f is what a slope measures. Everything below is assembled from slopes.

The divided difference of f at two points is the slope of the chord joining them, and at a repeated point it is the slope of the tangent.

Definition 21.9.1. Divided Difference

Let $f: I \rightarrow \mathbb{R}$ be differentiable. For $s, t \in I$, the **divided difference** of f is

$$f[s, t] := \begin{cases} \frac{f(s) - f(t)}{s - t}, & s \neq t, \\ f'(t), & s = t. \end{cases}$$

In words: for two distinct points it is the slope of the chord through $(s, f(s))$ and $(t, f(t))$; the second clause is the value that makes the first one continuous as s approaches t , which is exactly what the derivative is. It is symmetric, $f[s, t] = f[t, s]$, since swapping s and t negates both numerator and denominator, and it is additive in f : $(f + g)[s, t] = f[s, t] + g[s, t]$ and $(cf)[s, t] = c f[s, t]$.

Examples.

- $f(t) = \alpha + \beta t$: then $f[s, t] = \beta$ for all s, t , including $s = t$. The divided difference of an affine function is constant — it has one slope.
- $f(t) = t^2$: then $f[s, t] = (s^2 - t^2)/(s - t) = s + t$ for $s \neq t$, and $f[t, t] = 2t = t + t$. So $f[s, t] = s + t$ throughout.
- $f(t) = t^k$ for $k \geq 1$: dividing $s^k - t^k$ by $s - t$ gives $f[s, t] = \sum_{p=0}^{k-1} s^p t^{k-1-p}$, which at $s = t$ reads $kt^{k-1} = f'(t)$, so again one formula covers both clauses.
- $f(t) = -1/t$ on $(0, \infty)$: then $-1/s + 1/t = (s - t)/(st)$, so $f[s, t] = 1/(st)$, and at $s = t$ this is $1/t^2 = f'(t)$.
- $f(t) = \sqrt{t}$ on $(0, \infty)$: rationalizing, $(\sqrt{s} - \sqrt{t})/(s - t) = 1/(\sqrt{s} + \sqrt{t})$, and at $s = t$ this is $1/(2\sqrt{t}) = f'(t)$.

Nonexample by minimal change. Drop differentiability and keep continuity: for $f(t) = |t|$ on $\mathbb{R}$ the first clause still makes sense, but $f[0, 0]$ has no value, since the chord slopes $f[s, 0] = \operatorname{sgn}(s)$ do not settle down as $s \rightarrow 0$. Differentiability fills in the diagonal; continuous differentiability makes the filled-in function continuous.

Lemma 21.9.1. Divided Differences Are Continuous

Let $f \in C^1(I)$. If (s_k) and (t_k) are sequences in I with $s_k \rightarrow s_0 \in I$ and $t_k \rightarrow t_0 \in I$, then $f[s_k, t_k] \rightarrow f[s_0, t_0]$. In particular, for each fixed $\alpha \in I$ the function $t \mapsto f[\alpha, t]$ is continuous on I .

Proof. Case 1: $s_0 \neq t_0$. For large k we have $s_k \neq t_k$, since $s_k - t_k \rightarrow s_0 - t_0 \neq 0$. So $f[s_k, t_k] = (f(s_k) - f(t_k))/(s_k - t_k)$ for large k , and the algebra of limits applies, the denominator having nonzero limit: the quotient tends to $(f(s_0) - f(t_0))/(s_0 - t_0) = f[s_0, t_0]$, using that f is continuous because it is differentiable.

Case 2: $s_0 = t_0$. Let $\varepsilon > 0$. Since f' is continuous at t_0 , there is $\delta > 0$ with $|f'(c) - f'(t_0)| < \varepsilon$ for every $c \in I$ with $|c - t_0| < \delta$. Take k large enough that $|s_k - t_0| < \delta$ and $|t_k - t_0| < \delta$. If $s_k = t_k$ then $f[s_k, t_k] = f'(s_k)$, which is within ε of $f'(t_0)$. If $s_k \neq t_k$, then by the mean value theorem (A6), applied to f on the interval with endpoints s_k and t_k — an interval contained in I — there is c_k strictly between them with

$$f[s_k, t_k] = \frac{f(s_k) - f(t_k)}{s_k - t_k} = f'(c_k).$$

Then $|c_k - t_0| < \delta$, so again $|f[s_k, t_k] - f'(t_0)| < \varepsilon$. Since $f[t_0, t_0] = f'(t_0)$, this proves the claim. ■

21.9.2 The Loewner matrix

Collect the slopes between all pairs drawn from a list of points, and a matrix appears.

The Loewner matrix of f at n points is the table of all the chord slopes between them.

Definition 21.9.2. Loewner Matrix

Let $f: I \rightarrow \mathbb{R}$ be differentiable and let $t_1, \dots, t_n \in I$. The **Loewner matrix** of f at these points is

$$L_f(t_1, \dots, t_n) := (f[t_i, t_j])_{i,j=1}^n \in M_n(\mathbb{R}).$$

Every entry is real and the matrix is **symmetric**, because the divided difference is; its diagonal entries are $f'(t_1), \dots, f'(t_n)$. A real symmetric matrix is Hermitian, so it makes sense to ask whether it is positive semidefinite — the only question this section asks about it. The points need not be distinct; if $t_i = t_j$ then rows i and j coincide.

Examples. All four use the divided differences computed above.

- $f(t) = \alpha + \beta t$: every entry is β , so $L_f = \beta \mathbf{J}$, where $\mathbf{J}$ is the all-ones matrix. Since $\mathbf{J} = \mathbf{1}\mathbf{1}^\top$ is positive semidefinite, $L_f \succeq \mathbf{O}$ exactly when $\beta \geq 0$ — which is precisely when the affine

function is operator monotone, by Proposition 17.11.2 one way and Proposition 17.11.1 the other.

- $f(t) = -1/t$ on $(0, \infty)$: the (i, j) entry is $1/(t_i t_j)$, so $L_f = \mathbf{d}\mathbf{d}^\top$ with $\mathbf{d} = (1/t_1, \dots, 1/t_n)$. A matrix of the form $\mathbf{d}\mathbf{d}^\top$ is positive semidefinite, so **every** Loewner matrix of $-1/t$ is $\succeq \mathbf{O}$, at every n . Compare Proposition 17.11.3.
- $f(t) = -1/(t+c)$ with $c \geq 0$: the same computation with $t_i + c$ in place of t_i gives the rank-one matrix with entries $1/((t_i + c)(t_j + c))$, again $\succeq \mathbf{O}$. Compare Corollary 17.11.5.
- $f(t) = \sqrt{t}$ on $(0, \infty)$: the (i, j) entry is $1/(\sqrt{t_i} + \sqrt{t_j})$. A matrix of the form $(1/(x_i + x_j))$ with $x_i > 0$ is called a **Cauchy matrix**.

Nonexample by minimal change. In the first example take $\alpha = 0$ and $\beta = 1$, so that $f(t) = t$ and $L_f = \mathbf{J} \succeq \mathbf{O}$. Raise the exponent from 1 to 2, and positivity is lost at once.

Example. The Loewner Matrix of the Square

Let $f(t) = t^2$ on $[0, \infty)$ and let $s \neq t$ be points of $[0, \infty)$. Compute $L_f(s, t)$ and decide whether it is positive semidefinite.

Solution. Since $f[s, t] = s + t$ and $f'(t) = 2t$,

$$L_f(s, t) = \begin{pmatrix} 2s & s+t \\ s+t & 2t \end{pmatrix}, \quad \det L_f(s, t) = 4st - (s+t)^2 = -(s-t)^2.$$

The determinant is a principal minor, and it is < 0 because $s \neq t$. By Theorem 13.1.4 (e), $L_f(s, t)$ is not positive semidefinite — it is indefinite, its determinant being negative. So the square already fails at $n = 2$, at **every** pair of distinct points, which is the numerical shadow of Example (Squaring does not preserve the Loewner order).

Example. A Cauchy Matrix

Let $f(t) = \sqrt{t}$ on $(0, \infty)$ and take the points 1, 4, 9. Compute $L_f(1, 4, 9)$ and decide whether it is positive semidefinite.

Solution. The square roots are 1, 2, 3, and the (i, j) entry is $1/(\sqrt{t_i} + \sqrt{t_j})$, so

$$L_f(1, 4, 9) = \begin{pmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \end{pmatrix}.$$

The first two leading principal minors are

$$\frac{1}{2} \quad \text{and} \quad \frac{1}{2} \cdot \frac{1}{4} - \left(\frac{1}{3}\right)^2 = \frac{1}{8} - \frac{1}{9} = \frac{1}{72}.$$

For the third, expand along the first row. The three minors are

$$\frac{1}{24} - \frac{1}{25} = \frac{1}{600}, \quad \frac{1}{18} - \frac{1}{20} = \frac{1}{180}, \quad \frac{1}{15} - \frac{1}{16} = \frac{1}{240},$$

so, with the alternating signs,

$$\det L_f(1, 4, 9) = \frac{1}{2} \cdot \frac{1}{600} - \frac{1}{3} \cdot \frac{1}{180} + \frac{1}{4} \cdot \frac{1}{240} = \frac{1}{1200} - \frac{1}{540} + \frac{1}{960}.$$

Over the common denominator 43200 the three terms are 36, 80, and 45, and $36 - 80 + 45 = 1$, so $\det L_f(1, 4, 9) = 1/43200$. All three leading principal minors are positive, so $L_f(1, 4, 9) \succ \mathbf{O}$ by Sylvester's criterion (Theorem 13.1.5 (d)), and in particular $\succeq \mathbf{O}$.

The contrast between the two examples is the point of the section: $\sqrt{t}$ is operator monotone and t^2 is not, and the Loewner matrices see the difference with no matrix inequality tested. The theorem below says this is no coincidence.

Warning. One size at a time. A Loewner matrix is built from **one** choice of n points. Positivity for one choice says nothing about another, and positivity at all choices of n points is not by itself a statement about $n+1$ points: $f(t) = t^2$ has $L_f(t) = (2t) \succeq \mathbf{O}$ at every point of $[0, \infty)$, and already at two points it fails ([Example \(The Loewner Matrix of the Square\)](#)). The criterion below is an equivalence *size by size*: it matches the $n \times n$ Loewner matrices with order-preservation on $n \times n$ matrices, and operator monotonicity is the conjunction of all of these.

Since the criterion works one size at a time, the property it characterizes needs a name.

Definition 21.9.3. Monotone of Order n

Let $n \geq 1$ and let $f: I \rightarrow \mathbb{R}$. Then f is **monotone of order n on I** if, for all Hermitian $\mathbf{A}, \mathbf{B} \in M_n(F)$ whose spectra lie in I ,

$$\mathbf{A} \succeq \mathbf{B} \implies f(\mathbf{A}) \succeq f(\mathbf{B}).$$

Comparing with Definition 17.11.1 (a): f is **operator monotone on I exactly when it is monotone of order n for every $n \geq 1$** . Order 1 is plain monotonicity (Proposition 17.11.1), and order 2 is already a strictly stronger demand: $t \mapsto t^2$ is increasing on $[0, \infty)$ and fails at $n = 2$ ([Example \(Squaring does not preserve the Loewner order\)](#)). Whether each further order is in turn strictly stronger than the one before is not decided here.

21.9.3 Differentiating a matrix function

The link between the Loewner matrix and the order is a derivative, and the derivative comes out of one exact algebraic identity. Write both matrices in their spectral resolutions, insert the identity on both sides, and the difference of the two functional calculi factors through the difference of the matrices, with divided differences as the coefficients.

Lemma 21.9.2. The Two-Resolution Identity

Let $f: I \rightarrow \mathbb{R}$ be differentiable and let $\mathbf{A}, \mathbf{B} \in M_n(F)$ be Hermitian with spectra in I . Write the spectral resolutions

$$\mathbf{A} = \sum_i \alpha_i \mathbf{P}_i, \quad \mathbf{B} = \sum_k \beta_k \mathbf{Q}_k$$

with α_i the distinct eigenvalues of $\mathbf{A}$ and β_k those of $\mathbf{B}$ (Theorem 12.8.1). Then

$$f(\mathbf{B}) - f(\mathbf{A}) = \sum_{i,k} f[\alpha_i, \beta_k] \mathbf{P}_i (\mathbf{B} - \mathbf{A}) \mathbf{Q}_k.$$

Idea. The resolutions of the identity $\sum_i \mathbf{P}_i = \mathbf{I}$ and $\sum_k \mathbf{Q}_k = \mathbf{I}$ can be inserted on the two sides of any matrix, which splits it into $i \times k$ pieces. On the piece $\mathbf{P}_i \cdot \mathbf{Q}_k$, the matrix $\mathbf{A}$ acts as the number α_i from the left and $\mathbf{B}$ acts as β_k from the right. So both $f(\mathbf{B}) - f(\mathbf{A})$ and $\mathbf{B} - \mathbf{A}$ become numbers on each piece, namely $f(\beta_k) - f(\alpha_i)$ and $\beta_k - \alpha_i$, and the ratio of those two numbers is the divided difference.

Proof. By Theorem 12.8.1 (c), $\sum_i \mathbf{P}_i = \mathbf{I}$ and $\sum_k \mathbf{Q}_k = \mathbf{I}$. By Theorem 12.8.1 (e), $\mathbf{P}_i \mathbf{A} = \alpha_i \mathbf{P}_i$ and $\mathbf{B} \mathbf{Q}_k = \beta_k \mathbf{Q}_k$; and $\mathbf{P}_i f(\mathbf{A}) = f(\alpha_i) \mathbf{P}_i$, $f(\mathbf{B}) \mathbf{Q}_k = f(\beta_k) \mathbf{Q}_k$ follow from Definition 12.8.2 together with Theorem 12.8.1 (b). Hence

$$\begin{aligned} f(\mathbf{B}) - f(\mathbf{A}) &= \sum_{i,k} \mathbf{P}_i (f(\mathbf{B}) - f(\mathbf{A})) \mathbf{Q}_k \\ &= \sum_{i,k} (f(\beta_k) - f(\alpha_i)) \mathbf{P}_i \mathbf{Q}_k, \end{aligned}$$

and in the same way

$$\mathbf{P}_i (\mathbf{B} - \mathbf{A}) \mathbf{Q}_k = (\beta_k - \alpha_i) \mathbf{P}_i \mathbf{Q}_k.$$

Fix i and k . If $\alpha_i \neq \beta_k$, then by Definition 21.9.1

$$(f(\beta_k) - f(\alpha_i)) \mathbf{P}_i \mathbf{Q}_k = f[\alpha_i, \beta_k] (\beta_k - \alpha_i) \mathbf{P}_i \mathbf{Q}_k = f[\alpha_i, \beta_k] \mathbf{P}_i (\mathbf{B} - \mathbf{A}) \mathbf{Q}_k.$$

If $\alpha_i = \beta_k$, then both $f(\beta_k) - f(\alpha_i)$ and $\beta_k - \alpha_i$ are 0, so both sides of that equation are the zero matrix and it holds as well. Summing over i and k gives the identity. ■

The identity is exact, with no limit in it. To turn it into a derivative we let $\mathbf{B}$ approach $\mathbf{A}$, which needs the functional calculus to be continuous. It is, and the proof shows what a mixed pair of eigenbases is for.

Lemma 21.9.3. Continuity of the Functional Calculus

Let $J = [a, b]$ be a closed bounded interval, let $g: J \rightarrow \mathbb{R}$ be continuous, and let $n \geq 1$. For every $\varepsilon > 0$ there is $\eta > 0$ such that all Hermitian $\mathbf{X}, \mathbf{Y} \in M_n(F)$ with spectra in J and $\|\mathbf{X} - \mathbf{Y}\|_F < \eta$ satisfy

$$\|g(\mathbf{X}) - g(\mathbf{Y})\|_F < \varepsilon.$$

Idea. Diagonalize $\mathbf{X}$ and $\mathbf{Y}$ separately, and read both $g(\mathbf{X}) - g(\mathbf{Y})$ and $\mathbf{X} - \mathbf{Y}$ in the **mixed** basis, one eigenbasis on the left and the other on the right. In that basis the two matrices have entries $(g(a_i) - g(b_k))w_{ik}$ and $(a_i - b_k)w_{ik}$ with the same w_{ik} , so they differ only in how the scalar factor is built. Where a_i and b_k are close, continuity of g makes the first factor small; where they are far apart, the second factor is large and can pay for the first. The Frobenius norm is the right yardstick because it is computed from the entries and is

unchanged by a unitary factor on either side.

Proof. By the extreme value theorem (A4), g is bounded on the compact interval J ; let $M = \max_{t \in J} |g(t)|$. If $M = 0$, then g is identically 0 and there is nothing to prove, so assume $M > 0$.

Step 1: g is uniformly continuous on J . Suppose not. Then there is $\varepsilon_0 > 0$ such that for every $k \geq 1$ there are $a_k, b_k \in J$ with $|a_k - b_k| \leq 1/k$ and $|g(a_k) - g(b_k)| \geq \varepsilon_0$. By compactness (A3), a subsequence a_{k_j} converges to some $a \in J$; then $b_{k_j} \rightarrow a$ too, since $|b_{k_j} - a_{k_j}| \rightarrow 0$. Continuity of g at a gives $g(a_{k_j}) \rightarrow g(a)$ and $g(b_{k_j}) \rightarrow g(a)$, so $|g(a_{k_j}) - g(b_{k_j})| \rightarrow 0$, contradicting the lower bound ε_0 .

Step 2: the mixed basis. Let $\mathbf{X}$ and $\mathbf{Y}$ be Hermitian with spectra in J . By Corollary 12.4.3 over $\mathbb{C}$, or Corollary 12.5.2 over $\mathbb{R}$, there are unitary $\mathbf{U}, \mathbf{V} \in M_n(F)$ whose columns $\mathbf{u}_1, \dots, \mathbf{u}_n$ and $\mathbf{v}_1, \dots, \mathbf{v}_n$ are orthonormal eigenbases, with $\mathbf{X}\mathbf{u}_i = a_i\mathbf{u}_i$ and $\mathbf{Y}\mathbf{v}_k = b_k\mathbf{v}_k$, all $a_i, b_k \in J$. Put $w_{ik} = \mathbf{u}_i^* \mathbf{v}_k$, the entries of the unitary matrix $\mathbf{W} = \mathbf{U}^* \mathbf{V}$; then $\sum_{i,k} |w_{ik}|^2 = \|\mathbf{W}\|_F^2 = n$.

Write $\mathbf{X} = \sum_j \mu_j \mathbf{P}_j$ for the spectral resolution. Exactly one μ_j equals a_i , and $\mathbf{u}_i$ lies in its eigenspace, so $\mathbf{P}_j \mathbf{u}_i = \mathbf{u}_i$ for that j and $\mathbf{P}_j \mathbf{u}_i = \mathbf{0}$ for the others, the eigenspaces being orthogonal (Theorem 12.8.1 (a)). Hence $g(\mathbf{X})\mathbf{u}_i = g(a_i)\mathbf{u}_i$ by Definition 12.8.2, and likewise $g(\mathbf{Y})\mathbf{v}_k = g(b_k)\mathbf{v}_k$; since $g(a_i)$ is real and $g(\mathbf{X})$ is Hermitian, also $\mathbf{u}_i^* g(\mathbf{X}) = g(a_i)\mathbf{u}_i^*$. Hence the (i, k) entry of $\mathbf{U}^*(g(\mathbf{X}) - g(\mathbf{Y}))\mathbf{V}$ is

$$\mathbf{u}_i^*(g(\mathbf{X}) - g(\mathbf{Y}))\mathbf{v}_k = (g(a_i) - g(b_k))w_{ik},$$

and the same computation with g replaced by the identity function gives $\mathbf{u}_i^*(\mathbf{X} - \mathbf{Y})\mathbf{v}_k = (a_i - b_k)w_{ik}$. By Lemma 11.7.5,

$$\begin{aligned} \|g(\mathbf{X}) - g(\mathbf{Y})\|_F^2 &= \sum_{i,k} |g(a_i) - g(b_k)|^2 |w_{ik}|^2, \\ \|\mathbf{X} - \mathbf{Y}\|_F^2 &= \sum_{i,k} |a_i - b_k|^2 |w_{ik}|^2. \end{aligned}$$

Step 3: the split. Let $\varepsilon > 0$. By Step 1 there is $\delta > 0$ with $|g(a) - g(b)| \leq \varepsilon/(2\sqrt{n})$ whenever $a, b \in J$ and $|a - b| < \delta$. Split the index pairs (i, k) into those with $|a_i - b_k| < \delta$ and those with $|a_i - b_k| \geq \delta$. The first group contributes at most

$$\frac{\varepsilon^2}{4n} \sum_{i,k} |w_{ik}|^2 = \frac{\varepsilon^2}{4n} \cdot n = \frac{\varepsilon^2}{4}.$$

In the second group $|g(a_i) - g(b_k)| \leq 2M \leq (2M/\delta)|a_i - b_k|$, so that group contributes at most $(2M/\delta)^2 \|\mathbf{X} - \mathbf{Y}\|_F^2$. Therefore

$$\|g(\mathbf{X}) - g(\mathbf{Y})\|_F^2 \leq \frac{\varepsilon^2}{4} + \left(\frac{2M}{\delta}\right)^2 \|\mathbf{X} - \mathbf{Y}\|_F^2.$$

Take $\eta = \varepsilon\delta/(4M) > 0$. If $\|\mathbf{X} - \mathbf{Y}\|_F < \eta$, the second term is less than $\varepsilon^2/4$, so $\|g(\mathbf{X}) - g(\mathbf{Y})\|_F^2 < \varepsilon^2/2 < \varepsilon^2$. This proves the lemma. ■

Now the theorem. With the two lemmas in hand it is bookkeeping.

Theorem 21.9.4. The Derivative of a Matrix Function

Let $f \in C^1(I)$, let $\mathbf{A}, \mathbf{H} \in M_n(F)$ be Hermitian, and suppose there are $s_0 > 0$ and a closed bounded interval $J \subseteq I$ with

$$\text{spec}(\mathbf{A} + s\mathbf{H}) \subseteq J \quad \text{for every real } s \text{ with } |s| < s_0.$$

Then the matrix-valued function $\mathbf{G}(s) := f(\mathbf{A} + s\mathbf{H})$, defined for $|s| < s_0$, is differentiable at $s = 0$ in the sense of Definition 16.6.1, and:

(a) If $\mathbf{A} = \text{diag}(t_1, \dots, t_n)$, then

$$\mathbf{G}'(0) = L_f(t_1, \dots, t_n) \circ \mathbf{H},$$

the Hadamard product of Definition 13.8.1.

(b) In general, if $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^*$ with $\mathbf{U}$ unitary and $\mathbf{D} = \text{diag}(\lambda_1, \dots, \lambda_n)$, then

$$\mathbf{G}'(0) = \mathbf{U} \left(L_f(\lambda_1, \dots, \lambda_n) \circ (\mathbf{U}^* \mathbf{H} \mathbf{U}) \right) \mathbf{U}^*.$$

Remark. The hypothesis on J is automatic when I is **open**. Write $I = (c, d)$ with $-\infty \leq c < d \leq \infty$. Then $c < \lambda_n(\mathbf{A}) \leq \lambda_1(\mathbf{A}) < d$, so there is $\rho > 0$ with $[\lambda_n(\mathbf{A}) - \rho, \lambda_1(\mathbf{A}) + \rho] \subseteq I$; taking $s_0 = \rho/(\|\mathbf{H}\|_2 + 1)$ and J that closed interval works, because $|\lambda_i(\mathbf{A} + s\mathbf{H}) - \lambda_i(\mathbf{A})| \leq |s|\|\mathbf{H}\|_2 < \rho$ by Corollary 17.3.6.

Idea. Apply Lemma 21.9.2 with $\mathbf{B} = \mathbf{A} + s\mathbf{H}$. The right-hand side is a sum over pairs (eigenvalue of $\mathbf{A}$, eigenvalue of $\mathbf{B}$), and the sum over the second index is nothing but a functional calculus: freezing the first argument of the divided difference turns $f[\alpha_i, \cdot]$ into an ordinary continuous function g_i , and $\sum_k g_i(\beta_k) \mathbf{Q}_k$ is $g_i(\mathbf{B})$. So the identity collapses to

$$f(\mathbf{A} + s\mathbf{H}) - f(\mathbf{A}) = s \sum_i \mathbf{P}_i \mathbf{H} g_i(\mathbf{A} + s\mathbf{H}),$$

with the factor s already pulled out. Now let $s \rightarrow 0$: by Lemma 21.9.3, $g_i(\mathbf{A} + s\mathbf{H}) \rightarrow g_i(\mathbf{A})$, and the limit of the sum is the Hadamard product in disguise.

Proof. Write $\mathbf{A} = \sum_i \alpha_i \mathbf{P}_i$ for the spectral resolution, with $\alpha_1, \dots, \alpha_r$ the distinct eigenvalues of $\mathbf{A}$, and put $\mathbf{B}_s = \mathbf{A} + s\mathbf{H}$, so that $\mathbf{G}(s) = f(\mathbf{B}_s)$, for $|s| < s_0$. For each i define

$$g_i: J \rightarrow \mathbb{R}, \quad g_i(x) = f[\alpha_i, x],$$

which is continuous by Lemma 21.9.1, since $f \in C^1(I)$ and $J \subseteq I$.

Step 1: an exact formula. Let $\mathbf{B}_s = \sum_k \beta_k(s) \mathbf{Q}_k(s)$ be the spectral resolution of $\mathbf{B}_s$. By

Lemma 21.9.2 with $\mathbf{B} = \mathbf{B}_s$, and $\mathbf{B}_s - \mathbf{A} = s\mathbf{H}$,

$$\begin{aligned} f(\mathbf{B}_s) - f(\mathbf{A}) &= \sum_{i,k} f[\alpha_i, \beta_k(s)] \mathbf{P}_i(s\mathbf{H})\mathbf{Q}_k(s) \\ &= s \sum_i \mathbf{P}_i \mathbf{H} \left(\sum_k g_i(\beta_k(s)) \mathbf{Q}_k(s) \right) = s \sum_i \mathbf{P}_i \mathbf{H} g_i(\mathbf{B}_s), \end{aligned}$$

the last equality being Definition 12.8.2 applied to g_i and $\mathbf{B}_s$, whose spectrum lies in J , the domain of g_i .

Step 2: the limit. Since $\|\mathbf{B}_s - \mathbf{A}\|_F = |s| \|\mathbf{H}\|_F \rightarrow 0$ as $s \rightarrow 0$, and since $\mathbf{A}$ and every $\mathbf{B}_s$ are Hermitian with spectra in J , **Lemma 21.9.3** gives $\|g_i(\mathbf{B}_s) - g_i(\mathbf{A})\|_F \rightarrow 0$ for each i . By Definition 12.8.2,

$$g_i(\mathbf{A}) = \sum_j g_i(\alpha_j) \mathbf{P}_j = \sum_j f[\alpha_i, \alpha_j] \mathbf{P}_j.$$

Dividing the formula of Step 1 by $s \neq 0$ and letting $s \rightarrow 0$, the algebra of limits (applied entry by entry, and using $|m_{pq}| \leq \|\mathbf{M}\|_F$ for any matrix) gives

$$\frac{\mathbf{G}(s) - \mathbf{G}(0)}{s} \longrightarrow \sum_{i,j} f[\alpha_i, \alpha_j] \mathbf{P}_i \mathbf{H} \mathbf{P}_j.$$

Entrywise, this says exactly that each entry of $\mathbf{G}$ is differentiable at $s = 0$, with derivative the corresponding entry of the right-hand side. So $\mathbf{G}'(0)$ exists in the sense of Definition 16.6.1 and equals $\sum_{i,j} f[\alpha_i, \alpha_j] \mathbf{P}_i \mathbf{H} \mathbf{P}_j$.

Step 3: reading it as a Hadamard product. Let $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^*$ with $\mathbf{U}$ unitary, $\mathbf{D} = \text{diag}(\lambda_1, \dots, \lambda_n)$, and let $\mathbf{u}_1, \dots, \mathbf{u}_n$ be the columns of $\mathbf{U}$, so $\mathbf{A}\mathbf{u}_p = \lambda_p \mathbf{u}_p$. Each $\mathbf{P}_i$ is the orthogonal projection onto the eigenspace of α_i , which is spanned by those $\mathbf{u}_p$ with $\lambda_p = \alpha_i$; hence $\mathbf{P}_i = \sum_{p:\lambda_p=\alpha_i} \mathbf{u}_p \mathbf{u}_p^*$, and these sums partition $\{1, \dots, n\}$. Therefore

$$\sum_{i,j} f[\alpha_i, \alpha_j] \mathbf{P}_i \mathbf{H} \mathbf{P}_j = \sum_{p,q} f[\lambda_p, \lambda_q] \mathbf{u}_p \mathbf{u}_p^* \mathbf{H} \mathbf{u}_q \mathbf{u}_q^*.$$

Now $\mathbf{u}_p^* \mathbf{H} \mathbf{u}_q = (\mathbf{U}^* \mathbf{H} \mathbf{U})_{pq}$, and for any $\mathbf{M} \in M_n(F)$ we have $\mathbf{U} \mathbf{M} \mathbf{U}^* = \sum_{p,q} m_{pq} \mathbf{u}_p \mathbf{u}_q^*$. Taking $\mathbf{M} = L_f(\lambda_1, \dots, \lambda_n) \circ (\mathbf{U}^* \mathbf{H} \mathbf{U})$, whose (p, q) entry is $f[\lambda_p, \lambda_q] (\mathbf{U}^* \mathbf{H} \mathbf{U})_{pq}$, the displayed sum is $\mathbf{U} \mathbf{M} \mathbf{U}^*$. This proves (b), and (a) is the case $\mathbf{U} = \mathbf{I}$. ■

Remark. Two sanity checks. For $f(t) = t$ the Loewner matrix is $\mathbf{J}$, the identity for $\circ$, and the formula returns $\mathbf{H}$ — as it must, since $f(\mathbf{A} + s\mathbf{H}) = \mathbf{A} + s\mathbf{H}$. And the diagonal entries of (a) are $f'(t_i)h_{ii}$: to first order the i -th eigenvalue of $\text{diag}(t_1, \dots, t_n) + s\mathbf{H}$ moves by h_{ii} when the t_i are distinct (Theorem 20.11.2, as computed in Chapter 20 §11), and the chain rule then moves f of it by $f'(t_i)h_{ii}$. What the theorem adds is the **off-diagonal** part, where the eigenvectors turn, and there the chord slope $f[t_i, t_j]$ appears in place of the tangent slope.

Quick check. Take $n = 2$, $\mathbf{A} = \text{diag}(t_1, t_2)$ with $t_1 \neq t_2$, and $\mathbf{H} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. What does **Theorem 21.9.4** (a) predict, and what does it say about the eigenvalues of $\mathbf{A} + s\mathbf{H}$ to first order?

Solution. The Hadamard product with $\mathbf{H}$ keeps only the off-diagonal entries, so the derivative is $\begin{pmatrix} 0 & f[t_1, t_2] \\ f[t_1, t_2] & 0 \end{pmatrix}$. Its diagonal is zero, which matches the eigenvalue statement: $h_{11} = h_{22} = 0$, so to first order neither eigenvalue of $\mathbf{A} + s\mathbf{H}$ moves. All of the first-order change in $f(\mathbf{A} + s\mathbf{H})$ is in the eigenvectors, and the coefficient that measures it is the chord slope between t_1 and t_2 , not a derivative.

21.9.4 The criterion

Everything is now in place. One direction walks along the segment from $\mathbf{B}$ to $\mathbf{A}$, using the Schur product theorem to keep the derivative positive semidefinite; the other reads the derivative off in a single well-chosen direction.

Theorem 21.9.5. Loewner's Matrix Criterion

Let I be an interval with more than one point, let $f \in C^1(I)$ and let $n \geq 1$. The following are equivalent.

- (i) f is monotone of order n on I .
- (ii) $L_f(t_1, \dots, t_n) \succeq \mathbf{O}$ for all $t_1, \dots, t_n \in I$.

Idea. For (ii) $\Rightarrow$ (i): walk from $\mathbf{B}$ to $\mathbf{A}$ along the segment $\mathbf{B}_s = \mathbf{B} + s(\mathbf{A} - \mathbf{B})$, and watch the number $\varphi(s) = \langle f(\mathbf{B}_s)\mathbf{x}, \mathbf{x} \rangle$ for a fixed $\mathbf{x}$. Its derivative is, in the eigenbasis of $\mathbf{B}_s$, the quadratic form of a Hadamard product of two positive semidefinite matrices: the Loewner matrix, positive by hypothesis, and the direction $\mathbf{A} - \mathbf{B}$, positive because $\mathbf{A} \succeq \mathbf{B}$. The Schur product theorem says that such a product is positive semidefinite, so $\varphi' \geq 0$, and the mean value theorem does the rest.

For (i) $\Rightarrow$ (ii): the derivative formula turns order-preservation into $L_f \circ \mathbf{H} \succeq \mathbf{O}$ for every $\mathbf{H} \succeq \mathbf{O}$. Choose the one $\mathbf{H}$ that leaves L_f alone, namely the all-ones matrix $\mathbf{J} = \mathbf{1}\mathbf{1}^\top$, which is positive semidefinite and is the identity for $\circ$.

Proof. (ii) $\Rightarrow$ (i). Let $\mathbf{A}, \mathbf{B} \in M_n(F)$ be Hermitian with spectra in I and $\mathbf{A} \succeq \mathbf{B}$. Put $\mathbf{H} = \mathbf{A} - \mathbf{B} \succeq \mathbf{O}$ and $\mathbf{B}_s = \mathbf{B} + s\mathbf{H}$ for $s \in [0, 1]$. For such s , both $\mathbf{A} - \mathbf{B}_s = (1 - s)\mathbf{H}$ and $\mathbf{B}_s - \mathbf{B} = s\mathbf{H}$ are $\succeq \mathbf{O}$, so $\mathbf{A} \succeq \mathbf{B}_s \succeq \mathbf{B}$ and Corollary 17.2.4 gives

$$\lambda_i(\mathbf{B}) \leq \lambda_i(\mathbf{B}_s) \leq \lambda_i(\mathbf{A}) \quad (i = 1, \dots, n).$$

Hence $\text{spec}(\mathbf{B}_s) \subseteq J := [\lambda_n(\mathbf{B}), \lambda_1(\mathbf{A})]$ for every $s \in [0, 1]$, and $J \subseteq I$ because its two endpoints are eigenvalues of $\mathbf{B}$ and of $\mathbf{A}$, hence lie in I , and I is an interval.

Fix $\mathbf{x} \in F^n$ and put $\varphi(s) = \langle f(\mathbf{B}_s)\mathbf{x}, \mathbf{x} \rangle$ for $s \in [0, 1]$, a real number because $f(\mathbf{B}_s)$ is Hermitian (Theorem 12.8.3 (c)).

φ is continuous on $[0, 1]$. We have $\|\mathbf{B}_s - \mathbf{B}_{s'}\|_F = |s - s'| \|\mathbf{H}\|_F$, so Lemma 21.9.3 applied to f on J makes $s \mapsto f(\mathbf{B}_s)$ continuous in the Frobenius norm, and $|\langle \mathbf{M}\mathbf{x}, \mathbf{x} \rangle| \leq \|\mathbf{M}\|_F \|\mathbf{x}\|^2$ transfers this to φ .

$\varphi'(s) \geq 0$ for $0 < s < 1$. Fix such an s and put $s_0 = \min(s, 1 - s) > 0$. For $|\sigma| < s_0$ we have

$\mathbf{B}_s + \sigma \mathbf{H} = \mathbf{B}_{s+\sigma}$ with $s + \sigma \in (0, 1)$, so its spectrum lies in J . Thus [Theorem 21.9.4](#) applies to the pair $\mathbf{B}_s, \mathbf{H}$. Write $\mathbf{B}_s = \mathbf{U} \mathbf{D} \mathbf{U}^*$ with $\mathbf{U}$ unitary and $\mathbf{D} = \text{diag}(\lambda_1(\mathbf{B}_s), \dots, \lambda_n(\mathbf{B}_s))$, and set

$$\mathbf{C} = L_f(\lambda_1(\mathbf{B}_s), \dots, \lambda_n(\mathbf{B}_s)), \quad \mathbf{M} = \mathbf{U}^* \mathbf{H} \mathbf{U}.$$

Then $\mathbf{C} \succeq \mathbf{O}$ by hypothesis (ii), since the $\lambda_i(\mathbf{B}_s)$ lie in $J \subseteq I$, and $\mathbf{M} \succeq \mathbf{O}$ by Proposition 13.1.2 (a). By the Schur product theorem (Theorem 13.8.2 (a)), $\mathbf{C} \circ \mathbf{M} \succeq \mathbf{O}$. A quadratic form $\langle \mathbf{X} \mathbf{x}, \mathbf{x} \rangle$ is a fixed linear combination of the entries of $\mathbf{X}$, so an entrywise derivative of a matrix-valued function differentiates it; hence φ is differentiable at s with

$$\varphi'(s) = \langle \mathbf{U}(\mathbf{C} \circ \mathbf{M}) \mathbf{U}^* \mathbf{x}, \mathbf{x} \rangle = \langle (\mathbf{C} \circ \mathbf{M}) \mathbf{y}, \mathbf{y} \rangle \geq 0,$$

where $\mathbf{y} = \mathbf{U}^* \mathbf{x}$.

Conclusion. Let $0 < a < b < 1$. By the mean value theorem (A6) applied to φ on $(0, 1)$, there is c between a and b with $\varphi(b) - \varphi(a) = \varphi'(c)(b - a) \geq 0$. Letting $a \rightarrow 0^+$ and $b \rightarrow 1^-$ and using the continuity of φ on $[0, 1]$, we get $\varphi(1) - \varphi(0) \geq 0$, that is

$$\langle f(\mathbf{A}) \mathbf{x}, \mathbf{x} \rangle \geq \langle f(\mathbf{B}) \mathbf{x}, \mathbf{x} \rangle.$$

As $\mathbf{x}$ was arbitrary, $f(\mathbf{A}) \succeq f(\mathbf{B})$ by Definition 13.5.1.

(i) $\Rightarrow$ (ii). Suppose first that $t_1, \dots, t_n$ all lie in the interior of I , an open interval (c, d) . Put

$$\mathbf{A} = \text{diag}(t_1, \dots, t_n), \quad \mathbf{H} = \mathbf{J} = \mathbf{1} \mathbf{1}^\top,$$

so $\mathbf{H} \succeq \mathbf{O}$. Let $\rho > 0$ be small enough that $[\min_i t_i - \rho, \max_i t_i + \rho] \subseteq (c, d)$; this is possible because $c < \min_i t_i \leq \max_i t_i < d$. Put $s_0 = \rho / (\|\mathbf{H}\|_2 + 1)$ and $J = [\min_i t_i - \rho, \max_i t_i + \rho]$. By Corollary 17.3.6, $\text{spec}(\mathbf{A} + s\mathbf{H}) \subseteq J \subseteq I$ for $|s| < s_0$, so [Theorem 21.9.4](#) applies.

For $0 < s < s_0$ we have $(\mathbf{A} + s\mathbf{H}) - \mathbf{A} = s\mathbf{H} \succeq \mathbf{O}$, and both matrices have spectra in I , so (i) gives $f(\mathbf{A} + s\mathbf{H}) \succeq f(\mathbf{A})$, that is

$$\frac{f(\mathbf{A} + s\mathbf{H}) - f(\mathbf{A})}{s} \succeq \mathbf{O}.$$

For each fixed $\mathbf{x}$ the number $\langle \cdot, \mathbf{x} \rangle$ of the left-hand side is ≥ 0 , and it converges as $s \rightarrow 0^+$ to the quadratic form of the derivative, by [Theorem 21.9.4](#) (a); a nonstrict inequality survives a limit, so the derivative is $\succeq \mathbf{O}$. But the derivative is $L_f(t_1, \dots, t_n) \circ \mathbf{J} = L_f(t_1, \dots, t_n)$, because every entry of $\mathbf{J}$ is 1, so $\mathbf{J}$ is the identity for the Hadamard product (Definition 13.8.1). Hence $L_f(t_1, \dots, t_n) \succeq \mathbf{O}$.

Now let $t_1, \dots, t_n$ be arbitrary points of I . Since I has more than one point, it has an interior point m ; put $t_i^{(k)} = (1 - 1/k)t_i + (1/k)m$ for $k \geq 1$. Each $t_i^{(k)}$ lies in I , being a convex combination of two of its points, and it is an **interior** point: if $(m - \varepsilon, m + \varepsilon) \subseteq I$, then every point within ε/k of $t_i^{(k)}$ is of the form $(1 - 1/k)t_i + (1/k)(m + u)$ with $|u| < \varepsilon$, hence lies in I . Also $t_i^{(k)} \rightarrow t_i$. By the case just proved, $L_f(t_1^{(k)}, \dots, t_n^{(k)}) \succeq \mathbf{O}$ for every k , and by [Lemma 21.9.1](#) each entry converges to the corresponding entry of $L_f(t_1, \dots, t_n)$. For each $\mathbf{x}$, the quadratic forms are ≥ 0 and converge to the quadratic form of $L_f(t_1, \dots, t_n)$, which is therefore ≥ 0 . This proves the theorem. ■

Remark. The hypothesis $f \in C^1(I)$ enters in three places: it gives the diagonal entries $f'(t_i)$ of the Loewner matrix a meaning, it makes [Lemma 21.9.1](#) available, and through that lemma it supplies the continuous functions g_i that the derivative formula needs. Whether it can be dropped is not settled here: a monotone-of-order- n function is not obviously differentiable. The next section returns to that point for operator monotone functions, where the answer turns out to be yes, at the price of an imported theorem.

Taking $n = 1$ recovers Proposition 17.11.1: $L_f(t) = (f'(t))$, so monotone of order 1 means $f' \geq 0$, which for a C^1 function is the same as increasing, by the mean value theorem (A6). The first genuinely new case is $n = 2$, and there the criterion is two inequalities about numbers.

Corollary 21.9.6. The Criterion at Order Two

Let I be an interval with more than one point and let $f \in C^1(I)$. Then f is monotone of order 2 on I if and only if

$$f'(t) \geq 0 \text{ for all } t \in I, \quad \text{and} \quad f[s, t]^2 \leq f'(s)f'(t) \text{ for all } s, t \in I.$$

Proof. By [Theorem 21.9.5](#) with $n = 2$, monotonicity of order 2 is equivalent to $L_f(s, t) \succeq \mathbf{O}$ for all $s, t \in I$. The matrix

$$L_f(s, t) = \begin{pmatrix} f'(s) & f[s, t] \\ f[s, t] & f'(t) \end{pmatrix}$$

is real symmetric, and its principal minors are the two diagonal entries $f'(s), f'(t)$ and the determinant $f'(s)f'(t) - f[s, t]^2$. By Theorem 13.1.4 (e), $L_f(s, t) \succeq \mathbf{O}$ if and only if all three are ≥ 0 . Ranging over all $s, t \in I$, these are exactly the two displayed conditions; the case $s = t$ of the second is the tautology $f'(t)^2 \leq f'(t)^2$. ■

In words: **f increases, and the slope of every chord is at most the geometric mean of the slopes of the tangents at its ends.** Three of Chapter 17 §11's examples can now be re-derived, and its nonexample refuted, without touching a matrix.

- $f(t) = -1/t$ on $(0, \infty)$: $f'(t) = 1/t^2 \geq 0$, and $f[s, t] = 1/(st)$, so $f[s, t]^2 = 1/(s^2t^2) = f'(s)f'(t)$. The condition holds with **equality** everywhere, which is the numerical form of the fact that L_f has rank one.
- $f(t) = \sqrt{t}$ on $(0, \infty)$: $f'(t) = 1/(2\sqrt{t}) \geq 0$, and $f[s, t] = 1/(\sqrt{s} + \sqrt{t})$, so

$$f'(s)f'(t) - f[s, t]^2 = \frac{(\sqrt{s} + \sqrt{t})^2 - 4\sqrt{st}}{4\sqrt{st}(\sqrt{s} + \sqrt{t})^2} = \frac{(\sqrt{s} - \sqrt{t})^2}{4\sqrt{st}(\sqrt{s} + \sqrt{t})^2} \geq 0.$$

- $f(t) = \alpha + \beta t$: $f' = \beta$ and $f[s, t] = \beta$, so the conditions are $\beta \geq 0$ and $\beta^2 \leq \beta^2$.
- $f(t) = t^2$ on $[0, \infty)$: $f' = 2t \geq 0$, but $f[s, t]^2 - f'(s)f'(t) = (s + t)^2 - 4st = (s - t)^2 > 0$ for $s \neq t$. So t^2 is **not** monotone of order 2, and therefore not operator monotone. This is Example (Squaring does not preserve the Loewner order) again, now with the failure located at every pair of distinct points rather than at one pair of matrices.

Warning. **Order two is not “increasing and concave”.** The second condition of [Corollary 21.9.6](#) looks like a concavity statement and is not one, in either direction. The function $f(t) = t - t^2/2$ on $(0, 1)$ is increasing ($f' = 1 - t > 0$) and concave, yet

$$f'(s)f'(t) - f[s, t]^2 = (1-s)(1-t) - \left(1 - \frac{s+t}{2}\right)^2 = st - \left(\frac{s+t}{2}\right)^2,$$

which equals $-((s-t)/2)^2$, negative for $s \neq t$; so f is not monotone of order 2. In the other direction, $g(t) = -\log(1-t)$ on $(0, 1)$ is increasing and **convex** and *is* monotone of order 2. Writing $a = 1-s$ and $b = 1-t$, so that $s-t = b-a$, we have $g'(s)g'(t) = 1/(ab)$ and $g[s, t] = (\log b - \log a)/(b-a)$, so the second condition reads

$$(\log b - \log a)^2 \leq \frac{(b-a)^2}{ab},$$

an inequality about the logarithm alone; [Exercise \(C2: The logarithm at order two\)](#) proves it.

21.9.5 What this gives, and what it does not

The criterion is an equivalence, so it reads either way. Right to left it manufactures operator monotone functions: exhibit positive semidefinite Loewner matrices at every order and the function preserves the order at every size. Left to right it refutes them cheaply, since one bad pair of points kills order 2 and hence operator monotonicity.

What it does not give is the classification. The conditions “ $L_f(t_1, \dots, t_n) \succeq \mathbf{O}$ for all n and all points” are still infinitely many, and turning them into the integral representation of Theorem 17.11.9 is the analytic half of Loewner’s theorem — the business of the next section, and the one place where the chapter spends an imported theorem. What is proved here needs no analysis beyond three items of the chapter’s list: the mean value theorem (A6), which turns a nonnegative derivative into monotonicity along the segment; and, inside [Lemma 21.9.3](#), the extreme value theorem (A4), which bounds g , and the compactness of a closed bounded interval (A3), which makes g uniformly continuous.

21.9.6 Exercises

21.9.6.1 A. Check your understanding

Exercise. A1

- Define $f[s, t]$ for $s = t$, and say what hypothesis on f the definition needs.
- Write down $L_f(t_1, t_2, t_3)$ in terms of divided differences, and say why it is Hermitian.
- State [Theorem 21.9.5](#), and say which of its two directions uses the Schur product theorem.
- Decide whether the following is correct, and justify your answer: if $L_f(t_1, t_2) \succeq \mathbf{O}$ for all $t_1, t_2 \in I$, then f is operator monotone on I .
- What does [Theorem 21.9.4](#) give when $\mathbf{H}$ is diagonal?

Solution.

- (a) $f[t, t] = f'(t)$ (Definition 21.9.1); the definition needs f to be differentiable on I , and the continuity statement Lemma 21.9.1 needs f' continuous as well.
- (b) It is the 3×3 matrix whose (i, j) entry is $f[t_i, t_j]$, with diagonal $f'(t_1), f'(t_2), f'(t_3)$. It is real and symmetric, because f is real-valued and $f[s, t] = f[t, s]$; a real symmetric matrix is Hermitian.
- (c) For $f \in C^1(I)$ and $n \geq 1$: f is monotone of order n on I if and only if $L_f(t_1, \dots, t_n) \succeq \mathbf{O}$ for all $t_1, \dots, t_n \in I$. The Schur product theorem (Theorem 13.8.2) is used in the direction (ii) $\Rightarrow$ (i), to show that the derivative along the segment is positive semidefinite.
- (d) Incorrect. Positivity of all 2×2 Loewner matrices gives monotonicity of order 2 only (Corollary 21.9.6); operator monotonicity demands every order. The warning after Example (A Cauchy Matrix) makes the point, and Definition 17.11.1 (a) quantifies over all n .
- (e) If $\mathbf{H} = \text{diag}(h_1, \dots, h_n)$ and $\mathbf{A} = \text{diag}(t_1, \dots, t_n)$, then $L_f \circ \mathbf{H}$ is the diagonal matrix $\text{diag}(f'(t_1)h_1, \dots, f'(t_n)h_n)$: the two matrices commute, the problem is scalar in each coordinate, and the answer is the ordinary chain rule n times.

21.9.6.2 B. Practice

Exercise. B1: Loewner matrices by hand

Compute $L_f(1, 2)$ for each of the following, and determine whether it is positive semidefinite. Justify your answer.

- (i) $f(t) = 4 + 3t$ on $\mathbb{R}$.
- (ii) $f(t) = t^2$ on $(0, \infty)$.
- (iii) $f(t) = -1/t$ on $(0, \infty)$.
- (iv) $f(t) = \sqrt{t}$ on $(0, \infty)$, at the points 1 and 4 instead.

Solution.

- (i) Every entry is 3, so $L_f(1, 2) = 3\mathbf{J} = \begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix}$. Its principal minors are 3, 3 and 0, all ≥ 0 , so it is $\succeq \mathbf{O}$ (Theorem 13.1.4 (e)).
- (ii) $f'(t) = 2t$ and $f[1, 2] = 3$, so $L_f(1, 2) = \begin{pmatrix} 2 & 3 \\ 3 & 4 \end{pmatrix}$, with determinant $8 - 9 = -1 < 0$. Not positive semidefinite.
- (iii) $f'(t) = 1/t^2$ and $f[1, 2] = 1/2$, so $L_f(1, 2) = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1/4 \end{pmatrix}$, with determinant $\frac{1}{4} - \frac{1}{4} = 0$ and diagonal entries > 0 . Positive semidefinite, of rank one; indeed it is $\mathbf{d}\mathbf{d}^\top$ with $\mathbf{d} = (1, \frac{1}{2})$.
- (iv) $f'(t) = 1/(2\sqrt{t})$ and $f[1, 4] = 1/(1+2) = \frac{1}{3}$, so $L_f(1, 4) = \begin{pmatrix} 1/2 & 1/3 \\ 1/3 & 1/4 \end{pmatrix}$, with determinant $\frac{1}{8} - \frac{1}{9} = \frac{1}{72} > 0$ and positive diagonal. Positive definite, and in particular $\succeq \mathbf{O}$; it is the top-left corner of the matrix of Example (A Cauchy Matrix).

Exercise. B2: A derivative, two ways

Let $\mathbf{A} = \text{diag}(1, 4)$, let $\mathbf{H} = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$, and let $f(t) = t^2$ on $(0, \infty)$.

- (a) Compute $L_f(1, 4) \circ \mathbf{H}$.
- (b) Expand $(\mathbf{A} + s\mathbf{H})^2$ and differentiate at $s = 0$ directly. Check that the two answers agree.

Solution.

- (a) $L_f(1, 4) = \begin{pmatrix} 2 & 5 \\ 5 & 8 \end{pmatrix}$, so

$$L_f(1, 4) \circ \mathbf{H} = \begin{pmatrix} 2 \cdot 1 & 5 \cdot 2 \\ 5 \cdot 2 & 8 \cdot 3 \end{pmatrix} = \begin{pmatrix} 2 & 10 \\ 10 & 24 \end{pmatrix}.$$

- (b) $(\mathbf{A} + s\mathbf{H})^2 = \mathbf{A}^2 + s(\mathbf{A}\mathbf{H} + \mathbf{H}\mathbf{A}) + s^2\mathbf{H}^2$, so the derivative at $s = 0$ is $\mathbf{A}\mathbf{H} + \mathbf{H}\mathbf{A}$. Now

$$\mathbf{A}\mathbf{H} = \begin{pmatrix} 1 & 2 \\ 8 & 12 \end{pmatrix}, \quad \mathbf{H}\mathbf{A} = \begin{pmatrix} 1 & 8 \\ 2 & 12 \end{pmatrix}, \quad \mathbf{A}\mathbf{H} + \mathbf{H}\mathbf{A} = \begin{pmatrix} 2 & 10 \\ 10 & 24 \end{pmatrix},$$

which agrees with (a). In general $(\mathbf{A}\mathbf{H} + \mathbf{H}\mathbf{A})_{ij} = (t_i + t_j)h_{ij}$ for diagonal $\mathbf{A}$, and $t_i + t_j$ is exactly the divided difference of t^2 .

Exercise. B3: Refuting order two

For each function, use [Corollary 21.9.6](#) to decide whether it is monotone of order 2 on the interval given. Justify your answer.

- (i) $f(t) = t/(t+1)$ on $(0, \infty)$.
- (ii) $f(t) = t^2 - 6t$ on $(0, 1)$.
- (iii) $f(t) = 2t - 1/t$ on $(0, \infty)$.
- (iv) $f(t) = t + t^2$ on $(0, \infty)$.

Solution.

- (i) Monotone of order 2. Here $f(t) = 1 - 1/(t+1)$, so $f'(t) = 1/(t+1)^2 \geq 0$ and, by the computation for $-1/(t+c)$ with $c = 1$, $f[s, t] = 1/((s+1)(t+1))$. Hence $f[s, t]^2 = f'(s)f'(t)$, with equality. (In fact the Loewner matrices of every order are rank one and $\succeq \mathbf{O}$, which recovers [Corollary 17.11.5](#).)
- (ii) Not monotone of order 2 — indeed it fails already at order 1, since $f'(t) = 2t - 6 < 0$ on $(0, 1)$, so the first condition fails.
- (iii) Monotone of order 2. Here $f'(t) = 2 + 1/t^2 > 0$ and $f[s, t] = 2 + 1/(st)$, by additivity of divided differences and the two computations after [Definition 21.9.1](#). So

$$f'(s)f'(t) - f[s, t]^2 = \frac{2}{s^2} + \frac{2}{t^2} - \frac{4}{st} = 2\left(\frac{1}{s} - \frac{1}{t}\right)^2 \geq 0,$$

the terms 4 and $1/(s^2t^2)$ canceling between the two products.

- (iv) Not monotone of order 2. Here $f'(t) = 1 + 2t > 0$ and $f[s, t] = 1 + s + t$, so

$$\begin{aligned} f'(s)f'(t) - f[s, t]^2 &= (1 + 2s)(1 + 2t) - (1 + s + t)^2 \\ &= 4st - (s + t)^2 = -(s - t)^2, \end{aligned}$$

which is < 0 for $s \neq t$. Adding the operator monotone function t to the nonmonotone t^2 does not repair it.

21.9.6.3 C. Going deeper

Exercise. C1: Powers, directly

Let $k \geq 1$, let $\mathbf{A} = \text{diag}(t_1, \dots, t_n)$, and let $\mathbf{H} \in M_n(F)$ be Hermitian.

- Expand $(\mathbf{A} + s\mathbf{H})^k$ and show that the coefficient of s is $\sum_{p+q=k-1} \mathbf{A}^p \mathbf{H} \mathbf{A}^q$, the sum being over integers $p, q \geq 0$.
- Compute the (i, j) entry of that matrix and identify it with $f[t_i, t_j]h_{ij}$ for $f(t) = t^k$.
- Hence prove [Theorem 21.9.4](#) (a) for every polynomial f , without using [Lemma 21.9.3](#).

Solution.

- Expanding the product $(\mathbf{A} + s\mathbf{H})(\mathbf{A} + s\mathbf{H}) \cdots (\mathbf{A} + s\mathbf{H})$ of k factors and keeping the order of every product, the terms are indexed by the choice of $\mathbf{A}$ or $s\mathbf{H}$ in each factor. The terms with no $\mathbf{H}$ give $\mathbf{A}^k$; the terms with exactly one $\mathbf{H}$, in position $p + 1$, give $s \mathbf{A}^p \mathbf{H} \mathbf{A}^{k-1-p}$ for $p = 0, \dots, k - 1$; the remaining terms carry a factor s^m with $m \geq 2$. So

$$(\mathbf{A} + s\mathbf{H})^k = \mathbf{A}^k + s \sum_{p+q=k-1} \mathbf{A}^p \mathbf{H} \mathbf{A}^q + s^2 \mathbf{R}(s)$$

with $\mathbf{R}$ a matrix whose entries are polynomials in s .

- Since $\mathbf{A}^p = \text{diag}(t_1^p, \dots, t_n^p)$, the matrix $\mathbf{A}^p \mathbf{H} \mathbf{A}^q$ has (i, j) entry $t_i^p h_{ij} t_j^q$. Summing,

$$\left(\sum_{p+q=k-1} \mathbf{A}^p \mathbf{H} \mathbf{A}^q \right)_{ij} = h_{ij} \sum_{p=0}^{k-1} t_i^p t_j^{k-1-p} = h_{ij} f[t_i, t_j],$$

the last equality by the third example after [Definition 21.9.1](#), which covers $t_i = t_j$ as well.

- Each entry of $(\mathbf{A} + s\mathbf{H})^k$ is a polynomial in s , so it is differentiable at $s = 0$, and by (a) and (b) its derivative there is $f[t_i, t_j]h_{ij}$, that is, the (i, j) entry of $L_f(t_1, \dots, t_n) \circ \mathbf{H}$. By [Theorem 12.8.3](#) (a) the functional calculus of t^k is the k -th power, so this is the assertion for $f(t) = t^k$. For a general polynomial $f = \sum_k c_k t^k$, both $\mathbf{A} \mapsto f(\mathbf{A})$ and $f \mapsto L_f$ are linear in f ([Theorem 12.8.3](#) (b) and the additivity of divided differences), so the identity for the powers gives it for f .

Exercise. C2: The logarithm at order two

- Prove that $r - 1/r - 2 \log r \geq 0$ for every $r \geq 1$.
- Deduce that $\log$ is monotone of order 2 on $(0, \infty)$.

- (c) Deduce that $f(t) = -\log(1-t)$ is monotone of order 2 on $(0,1)$, and verify that f is convex there.

Hint for (a): differentiate, and recognize a square.

Solution.

- (a) Let $h(r) = r - 1/r - 2\log r$ on $[1, \infty)$. Then $h(1) = 0$ and, for $r > 0$,

$$h'(r) = 1 + \frac{1}{r^2} - \frac{2}{r} = \left(1 - \frac{1}{r}\right)^2 \geq 0.$$

For $r > 1$, the mean value theorem (A6) gives $h(r) - h(1) = h'(c)(r-1)$ for some $c \in (1, r)$, and both factors are ≥ 0 . Hence $h(r) \geq 0$.

- (b) By [Corollary 21.9.6](#) with $f = \log$ we must check $f'(t) = 1/t \geq 0$, which holds on $(0, \infty)$, and

$$\left(\frac{\log s - \log t}{s - t}\right)^2 \leq \frac{1}{st} \quad (s \neq t),$$

the case $s = t$ being the tautology. Both sides are unchanged by swapping s and t , so assume $s > t > 0$ and write $s = tr^2$ with $r > 1$. Then $\log s - \log t = 2\log r$, while

$$\frac{s - t}{\sqrt{st}} = \frac{t(r^2 - 1)}{tr} = r - \frac{1}{r}.$$

Both $2\log r$ and $r - 1/r$ are ≥ 0 , and (a) says $2\log r \leq r - 1/r$. Squaring the inequality between nonnegative numbers gives $(\log s - \log t)^2 \leq (s - t)^2/(st)$, which is the required inequality after dividing by $(s - t)^2 > 0$.

- (c) Put $a = 1 - s$ and $b = 1 - t$ for $s, t \in (0, 1)$, so $a, b \in (0, 1)$ and $s - t = b - a$. Then $f'(t) = 1/(1 - t)$, so $f'(s)f'(t) = 1/(ab)$, and

$$f[s, t] = \frac{-\log(1 - s) + \log(1 - t)}{s - t} = \frac{\log b - \log a}{b - a},$$

which is the divided difference of $\log$ at a and b . By (b) its square is at most $1/(ab) = f'(s)f'(t)$, and $f' > 0$, so [Corollary 21.9.6](#) applies. Finally $f''(t) = 1/(1 - t)^2 > 0$, so f is convex on $(0, 1)$ by Theorem 18.10.6. So an increasing **convex** function can be monotone of order 2.

Exercise. C3: Orders are nested

Let $1 \leq m \leq n$ and let $f: I \rightarrow \mathbb{R}$; in (b) and (c) assume in addition that $f \in C^1(I)$, so that the Loewner matrices are defined and [Theorem 21.9.5](#) applies.

- Prove that if f is monotone of order n on I , then it is monotone of order m on I .
- Give the corresponding statement about Loewner matrices, and prove it directly from [Definition 21.9.2](#), without using [Theorem 21.9.5](#).
- Explain why (a) and (b) together mean that a single failing pair of points rules out operator monotonicity at every order $n \geq 2$.

Hint for (a): pad with a scalar block.

Solution.

- (a) Let $\mathbf{A}, \mathbf{B} \in M_m(F)$ be Hermitian with spectra in I , $\mathbf{A} \succeq \mathbf{B}$. Fix $c \in I$ and put

$$\tilde{\mathbf{A}} = \mathbf{A} \oplus c\mathbf{I}_{n-m}, \quad \tilde{\mathbf{B}} = \mathbf{B} \oplus c\mathbf{I}_{n-m},$$

both Hermitian in $M_n(F)$ with spectra in I . Their difference is $(\mathbf{A} - \mathbf{B}) \oplus \mathbf{O}$, whose quadratic form at $(\mathbf{x}, \mathbf{y})$ is $\langle (\mathbf{A} - \mathbf{B})\mathbf{x}, \mathbf{x} \rangle \geq 0$, so $\tilde{\mathbf{A}} \succeq \tilde{\mathbf{B}}$. By hypothesis $f(\tilde{\mathbf{A}}) \succeq f(\tilde{\mathbf{B}})$. By Lemma 21.8.1 (c), $f(\tilde{\mathbf{A}}) = f(\mathbf{A}) \oplus f(c)\mathbf{I}_{n-m}$ and likewise for $\tilde{\mathbf{B}}$, so the difference is $(f(\mathbf{A}) - f(\mathbf{B})) \oplus \mathbf{O}$. Evaluating its quadratic form at $(\mathbf{x}, \mathbf{0})$ gives $\langle (f(\mathbf{A}) - f(\mathbf{B}))\mathbf{x}, \mathbf{x} \rangle \geq 0$ for every $\mathbf{x} \in F^m$, that is $f(\mathbf{A}) \succeq f(\mathbf{B})$.

- (b) If $L_f(t_1, \dots, t_n) \succeq \mathbf{O}$ for all $t_1, \dots, t_n \in I$, then $L_f(t_1, \dots, t_m) \succeq \mathbf{O}$ for all $t_1, \dots, t_m \in I$. Indeed, given $t_1, \dots, t_m$, pad the list with any $t_{m+1}, \dots, t_n \in I$; then $L_f(t_1, \dots, t_m)$ is the leading principal $m \times m$ submatrix of $L_f(t_1, \dots, t_n)$, because the (i, j) entry of each is $f[t_i, t_j]$ for $i, j \leq m$. A principal submatrix of a positive semidefinite matrix is positive semidefinite, since its quadratic form is the quadratic form of the big matrix at vectors supported on those coordinates (compare the proof of Theorem 13.1.4).
- (c) By (b), a pair $s, t \in I$ with $L_f(s, t) \not\succeq \mathbf{O}$ forces $L_f(t_1, \dots, t_n) \not\succeq \mathbf{O}$ for the padded list at every $n \geq 2$, so by Theorem 21.9.5 f is monotone of no order $n \geq 2$. Equivalently, by (a), failure at order 2 propagates upward — which is why Corollary 21.9.6 is the cheapest refutation available.