A Positivity and positive-definiteness

We recall that a real number $x \in \mathbb{R}$ is said to be (strictly) positive if $x > 0$, and nonnegative if $x \geq 0$. With a slight abuse of language, we shall say that a function $f : \mathbb{R}^d \to \mathbb{R}$ is positive (meaning "nonnegative") if $f(\boldsymbol{x}) \geq 0$, for all $\boldsymbol{x} \in \mathbb{R}^d$. Likewise, one extends the definition to tempered distributions.

Definition 27. *A tempered distribution $g : \mathcal{S}(\mathbb{R}^d) \to \mathbb{R}$ is said to be positive (also denoted as $g \geq 0$) if*

$$\langle g, \varphi \rangle \geq 0$$

for every positive test function $\varphi \in \mathcal{S}(\mathbb{R}^d)$ with $\varphi \geq 0$.

Positive distributions are also called *tempered measures* as they can always be represented in terms of a Lebesgue integral $\langle g, \varphi \rangle = \int_{\mathbb{R}^d} \varphi(\boldsymbol{x}) \mu_g(\mathrm{d}\boldsymbol{x})$ where the corresponding measure μ_g is allowed to exhibit some polynomial growth at infinity. In other words, there must exist some $k \in \mathbb{N}$ such that $\int_{\mathbb{R}^d} (1 + \|\boldsymbol{x}\|^2)^{-k} \mu_g(\mathrm{d}\boldsymbol{x}) < \infty$.

Positive-definiteness is a related property that identifies the important subclass of operators (or kernels) whose spectrum is positive. The notion comes in a variety of forms as it applies to symmetric matrices, kernels, linear operators, bilinear forms, functions, generalized functions (or linear functionals) as well as general functionals (not necessarily linear). These can all be unified if one starts from the following abstract definition.

Definition 28 (Positive-definiteness)**.** *Let E be some arbitrary set and h a bivariate complex-valued function $E \times E \to \mathbb{C}$. Then, h is said to be positive-semi-definite (*positive-definite*, for short) on $E \times E$ if*

$$\sum_{m=1}^{N} \sum_{n=1}^{N} \overline{z}_m h(x_m, x_n) z_n \geq 0 \tag{123}$$

for every possible choice of $x_1, \dots, x_N \in E$, $z_1, \dots, z_N \in \mathbb{C}$, and any positive integer N. Likewise, h is said to be strictly positive-definite *if* (123) *can be replaced by a strict inequality for any $\mathbf{z} \in \mathbb{C}^N \backslash \{0\}$ and any series of distinct points $x_n \in E$.*

The four special instances that are relevant for our purpose are:

1. A square-matrix matrix $\mathbf{H} \in \mathbb{C}^{N \times N}$, whose entries are indexed by $m, n \in E = \{1, 2, \dots, N\}$ and specify the map

$$h : (m, n) \mapsto h[m, n] = [\mathbf{H}]_{m,n}.$$

Hence, the positive-definiteness of $\mathbf{H}$ is equivalent to

$$\sum_{m=1}^{N}\sum_{n=1}^{N}\overline{z}_m h[m,n] z_n = \mathbf{z}^H\mathbf{H}\mathbf{z} \geq 0, \tag{124}$$

for any $\mathbf{z} = (z_1, \dots, z_N) \in \mathbb{C}^N$.

2. A bivariate kernel $h : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$, which is indexed by $\boldsymbol{x}, \boldsymbol{y} \in E = \mathbb{R}^d$. In this case, the above definition of positive-definiteness is equivalent to the standard one in Definition 2.

3. A function $f : \mathbb{R}^d \to \mathbb{C}$. The positive-definiteness of the latter (see Definition 40) is equivalent to the positive-definiteness of the "shift-invariant" kernel $h(\boldsymbol{x}, \boldsymbol{y}) = f(\boldsymbol{x} - \boldsymbol{y})$ with $\boldsymbol{x}, \boldsymbol{y} \in E = \mathbb{R}^d$.

4. A complex-valued functional $F : \mathcal{X} \to \mathbb{C}$ defined over some topological space $\mathcal{X} = E$. Here too, the positive-definiteness of F (see Definition 25) is compatible with the above definition if we take $h(x, y) = F(x-y)$ with $x, y \in E = \mathcal{X}$.

The next observation is that, in the case of a matrix, positive definiteness implies Hermitian symmetry, as well as the positivity of the eigenvalues (spectrum).

Proposition 18 (Positive-definite matrix)**.** *The square matrix $\mathbf{H} \in \mathbb{C}^{N\times N}$ is positive-definite (resp., strictly positive-definite) if and only if it is Hermitian-symmetric and all its eigenvalues are non-negative (resp., strictly positive).*

Proof. The requirement that the double sum in (124) must be real-valued is expressed as $\overline{\mathbf{z}}^T\mathbf{H}\mathbf{z} = \overline{\overline{\mathbf{z}}^T\mathbf{H}\mathbf{z}} = \overline{\mathbf{z}}^T\overline{\mathbf{H}}^T\mathbf{z}$ for all $\mathbf{z} \in \mathbb{C}^N$, which is equivalent to $\overline{\mathbf{H}}^T = \mathbf{H}$ (Hermitian symmetry). Since $\mathbf{H}$ is Hermitian, it admits an eigendecomposition $\mathbf{H} = \mathbf{U}\,\mathrm{diag}(\lambda_1, \dots, \lambda_N)\,\mathbf{U}^H$ where the eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$ are real-valued and the eigenvector matrix $\mathbf{U} = (\mathbf{u}_1, \dots, \mathbf{u}_N)$ is unitary—i.e., such that, $\mathbf{U}^H\mathbf{U} = \mathbf{I}$. Consequently, we can rewrite $\mathbf{z}^H\mathbf{H}\mathbf{z}$ as

$$\mathbf{z}^H\mathbf{H}\mathbf{z} = \sum_{n=1} \lambda_n |y_n|^2$$

where $y_n = \mathbf{u}_n^H\mathbf{z}$. Since the eigenvectors $\{\mathbf{u}_n\}_{n=1}^N$ form a basis of $\mathbb{C}^N$, the above quantity is nonnegative for all $\mathbf{z}$ iff. the eigenvalues λ_n are all nonnegative. Likewise, we deduce that the Hermitian matrix $\mathbf{H}$ is strictly positive-definite iff. $\lambda_n > 0$ for $n = 1, \dots, N$. $\square$

Hence, the positive definiteness of the bivariate function h in Definition 28 is equivalent to the positive definiteness of all "sampling" matrices $\mathbf{H}$ whose entries are given by $[\mathbf{H}]_{m,n} = h(x_m, x_n)$. Moreover, if the positive definiteness is strict, then the matrices $\mathbf{H}$ are all invertible. We now list some of the structural implications of this connection.

Proposition 19 (Hermitian symmetry)**.** *Let E be some set and h a bivariate complex-valued function $E \times E \to \mathbb{C}$. Then, h enjoys the following properties:*

- *Non-negative diagonal: $h(x,x) \geq 0$ for all $x \in E$*
- *Hermitian symmetry: $\overline{h(x,y)} = h(y,x)$ for all $x, y \in E$*
- $|h(x,y)| \leq \sqrt{h(x,x)}\sqrt{h(y,y)} \leq \frac{h(x,x)+h(y,y)}{2}$ *for all $x, y \in E$.*

Proof. The first statement follows directly from (123) with $N = 1$ and $x_1 = x$. To derive the two others, we take $N = 2$ with $x_1 = x$ and $x_2 = y$ and write the corresponding sampling matrix

$$\mathbf{H}_2 = \begin{pmatrix} h(x,x) & h(x,y) \\ h(y,x) & h(y,y) \end{pmatrix},$$

which must be (semi-)positive-definite and hence Hermitian-symmetric (by Proposition 18). The final inequality is deduced from the nonnegativity of the determinant: $\det(\mathbf{H}_2) = \lambda_1\lambda_2 = h(x,x)h(y,y) - |h(x,y)|^2 \geq 0$. □

A positive-definite function (Item 3 in the list) can also be seen as a special case of a positive-definite functional with $\mathcal{X} = \mathbb{R}^d$, whose properties are now examined in more detail.

Proposition 20 (Boundedness of positive-definite functionals)**.** *Let $F : \mathcal{X} \to \mathbb{C}$ be a positive-definite functional over some topological vector space $\mathcal{X}$. Then, F enjoys the following properties*

- *Boundedness: $0 \leq |F(\varphi)| \leq F(0)$ for all $\varphi \in \mathcal{X}$*
- *Hermitian symmetry: $\overline{F(\varphi)} = F(-\varphi)$ for all $\varphi \in \mathcal{X}$*
- $|F(\varphi) - F(\phi)| \leq 2\sqrt{F(0)}\sqrt{|F(0) - F(\varphi - \phi)|}$ *for all $\varphi, \phi \in \mathcal{X}$*

Proof. The first two properties are deduced using the same technique as in the proof of Proposition 19. As for the third one, it is derived from the positive-definiteness of the "sampling" matrix with $N = 3, \varphi_1 = 0, \varphi_2 = \varphi$, and $\varphi_3 = \phi$:

$$\mathbf{H}_3 = \begin{pmatrix} F(0) & F(\varphi) & F(\phi) \\ F(-\varphi) & F(0) & F(\phi-\varphi) \\ F(-\phi) & F(\varphi-\phi) & F(0) \end{pmatrix}$$

with $F(-\phi) = \overline{F(\phi)}$, $F(-\varphi) = \overline{F(\varphi)}$ and $F(\phi-\varphi) = \overline{F(\varphi-\phi)}$ (Hermitian symmetry). The non-negativity of the determinant then yields

$$\begin{aligned} 0 \le \det(\mathbf{H}_3) =& F(0)^3 - F(0)|F(\varphi-\phi)|^2 - F(\varphi)\big(F(0)\overline{F(\varphi)} - \overline{F(\varphi-\phi)F(\phi)}\big) + \\ & F(\phi)\big(\overline{F(\varphi)}F(\varphi-\phi) - F(0)\overline{F(\phi)}\big) \\ \le& \, 4F(0)^2|F(0) - F(\varphi-\phi)| - F(0)|F(\varphi) - F(\phi)|^2 \end{aligned}$$

where we have made use of the inequality $a^3 - ab^2 \ge 2a^2|a-b|$ for $|b| < a$. □

A remarkable consequence of the last property in Proposition 20 is that the continuity of F in the neighborhood of $\varphi = 0$ is sufficient to ensure its continuity everywhere.

The notion of positive-definiteness may also be extended to objects that are generalized functions. This extension relies on the equivalence between kernels and linear operators, as stated in Schwartz' kernel theorem (Theorem 33). The resulting criterion should be viewed as the infinite-dimensional counterpart of the condition $\overline{\mathbf{z}}^T\mathbf{H}\mathbf{z} = \langle \mathbf{H}\mathbf{z}, \overline{\mathbf{z}}\rangle \ge 0$ that characterizes a positive-definite matrix.

Definition 29 (Positive-definiteness for generalized functions)**.** *In order to cover all Hermitian-symmetric kernels, we are momentarily considering the Schwartz spaces $\mathcal{S}(\mathbb{R}^d)$ and $\mathcal{S}'(\mathbb{R}^d)$ to be complex-valued.*

1. *Let $h \in \mathcal{S}'(\mathbb{R}^d \times \mathbb{R}^d)$ be the (complex-valued) Schwartz kernel of the continuous operator $\mathrm{H} : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d)$. Then, h is said to be positive-definite if $\langle \mathrm{H}\varphi, \overline{\varphi}\rangle \ge 0$ for all $\varphi \in \mathcal{X}$, which may also be written symbolically as*

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \overline{\varphi(\boldsymbol{x})} h(\boldsymbol{x},\boldsymbol{y})\varphi(\boldsymbol{y})\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{y} \ge 0.$$

2. *The (complex-valued) generalized function $g : \mathcal{S}(\mathbb{R}^d) \to \mathbb{C}$ is said to be positive-definite if*

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \overline{\varphi(\boldsymbol{x})} g(\boldsymbol{x}-\boldsymbol{y})\varphi(\boldsymbol{y})\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{y} = \langle g * \varphi, \overline{\varphi}\rangle = \langle g, (\overline{\varphi}^\vee * \varphi)\rangle \ge 0$$

where $\varphi^\vee$ denotes the reversed version φ; i.e., $\varphi^\vee(\boldsymbol{x}) \stackrel{\triangle}{=} \varphi(-\boldsymbol{x})$ for any $\boldsymbol{x} \in \mathbb{R}^d$.

The above formulation of positive definiteness is an extension of Definition 28 for $E = \mathbb{R}^d$ where the two finite sums have been replaced by integrals (these actually represent duality products with respect to the variables $\boldsymbol{x}$ and $\boldsymbol{y}$). The two formulations can be shown to be equivalent when h and g are ordinary bivariate or univariate continuous functions; that is, when $h \in C_{\mathrm{b},\alpha}(\mathbb{R}^d \times \mathbb{R}^d)$ and $g \in C_{\mathrm{b},\alpha}(\mathbb{R}^d)$ for some $\alpha \in \mathbb{R}$ (see [**?**, Proposition 12.3, p. 90]).

We also note that the second condition in Definition 28 is a special instance of the first with $h(\boldsymbol{x}, \boldsymbol{y}) = g(\boldsymbol{x} - \boldsymbol{y})$, in which case the underlying operator $\mathrm{H} : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d)$ is linear shift-invariant with impulse response $g = \mathrm{T}\{\delta\} \in \mathcal{S}'(\mathbb{R}^d)$. An interesting twist is that the latter convolutional form, which is more general than the classical definition of a positive-definite function (see Definition 40 in Appendix E.3), actually facilitates the derivation of Bochner's theorem (Theorem 35), which states the equivalence between this property and Fourier-domain positivity.

Theorem 26 (Schwartz-Bochner)**.** *A generalized function $\hat{g} \in \mathcal{S}'(\mathbb{R}^d)$ is positive-definite if and only if it is the generalized Fourier transform of a positive distribution with corresponding tempered measure $\mu_g \geq 0$; that is,*

$$\langle \hat{g}, \varphi \rangle = \langle g, \hat{\varphi} \rangle = \int_{\mathbb{R}^d} \hat{\varphi}(\boldsymbol{x}) \mu_g(\mathrm{d}\boldsymbol{x})$$

where $\hat{\varphi}(\boldsymbol{\omega}) = \int_{\mathbb{R}^d} \varphi(\boldsymbol{x}) \mathrm{e}^{-\mathrm{j}\langle \boldsymbol{\omega}, \boldsymbol{x} \rangle} \mathrm{d}\boldsymbol{x}$ is the Fourier transform of the test function $\varphi \in \mathcal{S}(\mathbb{R}^d)$. Moreover, if $\hat{g}$ is continuous at the origin with $\hat{g}(0) = 1$, then it is the (ordinary) Fourier transform of a Borel measure $\mu_g \geq 0$ with $\mu_g(\mathbb{R}^d) = 1$.

We outline the proof of Theorem 26, which takes advantage of two basic properties: (1) the product of two test functions is a test function, and, (2) the Fourier transform maps a convolution into a product. Consequently, $\phi = \varphi^\vee * \varphi$ is a test function whose Fourier transform is $\hat{\phi} = |\hat{\varphi}|^2$ where $\hat{\varphi} = \mathcal{F}\{\varphi\} \in \mathcal{S}(\mathbb{R}^d)$ is the (ordinary) Fourier transform of φ. Next, we recall that a generalized function $\hat{g} \in \mathcal{S}'(\mathbb{R}^d)$ is the generalized Fourier transform of $g \in \mathcal{S}'(\mathbb{R}^d)$—that is, $\hat{g} = \mathcal{F}\{g\}$—iff.

$$\langle \hat{g}, \varphi \rangle = \langle g, \hat{\varphi} \rangle$$

for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$. For the particular case of $\phi \in \mathcal{S}(\mathbb{R}^d)$, this yields

$$0 \leq \langle \hat{g}, (\varphi^\vee * \varphi) \rangle = \langle \hat{g}, \phi \rangle = \langle g, \hat{\phi} \rangle = \langle g, |\hat{\varphi}|^2 \rangle,$$

which shows that the positive definiteness of $\hat{g}$ is equivalent to the multiplicative positivity of its generalized inverse Fourier transform g; i.e., $\langle g, |\hat{\varphi}|^2\rangle \geq 0$ for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$. The final technical step is to show that the latter property is equivalent to the positivity of g in the sense of Definition 27. This is achieved by proving that every positive function $\psi(\boldsymbol{x}) \geq 0$ in $\mathcal{S}(\mathbb{R}^d)$ can be written as the limit of a series of functions $\overline{\varphi_m(\boldsymbol{x})}\varphi_m(\boldsymbol{x})$ with $\varphi_m \in \mathcal{S}(\mathbb{R}^d)$ (see [**?**, p. 149-150]).

As for the last statement in Theorem 26, we use the property that the continuity of $\hat{g}$ at the origin implies its continuity everywhere (see third statement in Proposition 20 with $\mathcal{X} = \mathbb{R}^d$). Hence, $\hat{g}(\boldsymbol{\omega})$ is well-defined pointwise for all $\boldsymbol{\omega} \in \mathbb{R}^d$ with $|\hat{g}(\boldsymbol{\omega})| \leq \hat{g}(0) = \int_{\mathbb{R}^d} \mu_g(\mathrm{d}\boldsymbol{x}) = \mu_g(\mathbb{R}^d) = 1$. The latter bound ensures that the generalized Fourier transform of g (or μ_g) can be expressed as a classical Fourier-Lebesgue integral; that is, $\hat{g}(\boldsymbol{\omega}) = \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{j}\langle\boldsymbol{\omega},\boldsymbol{x}\rangle} \mu_g(\mathrm{d}\boldsymbol{x})$.

More generally, we note that the property of μ_g being a tempered (positive) measure is equivalent to the boundedness of the positive measure $\mu_f = (1+\|\cdot\|^2)^{-k}\mu_g$ for some suitable k. This yields the following equivalent statement of Theorem 26.

Corollary 7. *$\hat{g} = \mathcal{F}\{\mu_g\}$ is a positive-definite distribution iff. there exist an integer k and a continuous positive-definite function $\hat{f}$ such that $\hat{g} = (\mathrm{Id} - \Delta)^k \hat{f}$ where $\Delta = \sum_{n=1}^d \frac{\partial^2}{(\partial\omega_n)^2}$ is the Laplacian operator.*