Experimental Positive Definiteness of the Gaussian Kernel

Editors: List of editors' names

Abstract

We study positive definiteness of the Gaussian kernel $\exp(-\lambda d(x,y)^2)$ on two-dimensional cones. Based on numerical experiments, we conjecture that the set of parameters λ that make the kernel positive definite is related to the cone angle. We relate our experimental results to existing positivity and non-positivity theorems in the literature.

Keywords: kernel, Gaussian kernel, positive definite, curvature, metric space

1. Introduction

For a metric space (X,d), the *Gaussian kernel* is given by the formula:

$$k(x,y) = \exp(-\lambda\, d(x,y)^2) \quad \text{for} \quad x,y \in X \tag{1}$$

for some parameter $\lambda > 0$. The Gaussian kernel frequently appears in pure mathematics and real life applications and it is often important that k is *positive definite* (see Definition 2). It is known that for some spaces, the kernel is positive definite for all $\lambda > 0$, while for others, it is not positive definite for any $\lambda > 0$.

Open problem: For a connected metric space (X,d), how does the range of values λ for which the Gaussian kernel is positive depend on the geometry of X?

This is a more basic version of (Li, 2024, Conjecture/Question). If one drops the *connected* condition, the problem has a different flavour, and we present one well known folklore example of a discrete space in section B.

We do not solve the open problem, but we numerically study certain cone spaces Y_α depending on one positive parameter $\alpha > 0$. We present numerical evidence suggesting that for $\alpha \in (0, 2\pi)$, the Gaussian kernel on Y_α is positive for some values of λ but not for all of them, and that the range of λ that makes the kernel positive gets smaller as α gets smaller. This can be related to (a discrete notion of) the curvature of Y_α.

2. Existing work

The Gaussian kernel is widely used in Support Vector Machines (SVMs) Cortes and Vapnik (1995); Schölkopf and Smola (2002), Gaussian Processes Rasmussen and Williams (2006), kernel density estimation Silverman (1986), and many other methods.

Positivity is an important property for kernels (see e.g. Aronszajn (1950)) and checking positivity of kernels on various geometries is an area of intense research activity. The foundational result for the Gaussian kernel is Schönberg's article Schoenberg (1938a), where it was shown that the Gaussian kernel is positive definite on $\mathbb{R}^n$ for all $\lambda > 0$. This positivity is a rare property of $\mathbb{R}^n$: Feragen et al. (2015) proved that positive definiteness for all λ on geodesic spaces implies flatness. Da Costa et al. (2023) proved that the Gaussian kernel fails to be positive on the circle S^1 for all $\lambda > 0$, and Li (2024) extended this to all compact, non-simply-connected manifolds. We know of one example of a connected space on which the Gaussian kernel is known to be positive definite for some but not all values of λ, namely (Sra, 2016, Theorem 3.10).

Theory aside, numerous recent studies have explored Gaussian kernels in diverse applications such as 3D processing Song et al. (2024); Feng et al. (2025), neuroscience Yuan et al. (2016), and quantum learning Bishwas et al. (2020).

3. Background

3.1. Kernels and Positive Definiteness

In this section, we fix some notation for working with kernels, taken from the textbook Schölkopf and Smola (2002).

Definition 1 *A function $k : X \times X \to \mathbb{R}$ is a* kernel *on a set X if it is symmetric, i.e., $k(x, y) = k(y, x)$ for all $x, y \in X$.*

Definition 2 (Definition 2.5 in Schölkopf and Smola (2002)) *A kernel k on X is* positive definite *if for any finite set of points $\{x_1, \ldots, x_n\} \subset X$ and any non-zero vector $c \in \mathbb{R}^n$, the following inequality holds:*

$$\sum_{i=1}^{n} \sum_{j=1}^{n} c_i c_j k(x_i, x_j) \geq 0 \tag{2}$$

This is equivalent to asking that the Gram matrix K, with entries $K_{ij} = k(x_i, x_j)$, is a positive semi-definite matrix.

3.2. Results on the Gaussian Kernel

The following three theorems give some information about when the Gaussian kernel is positive definite:

Theorem 3 (Proposition 2.28 in Schölkopf and Smola (2002), cf. Schoenberg (1938b)) *On Euclidean space $\mathbb{R}^n$, with the standard Euclidean distance, the Gaussian kernel is positive definite for all $\lambda > 0$.*

Theorem 4 (Theorem 2 in Feragen et al. (2015)) *If (M, g) is a Riemannian manifold with geodesic distance d_g, and the Gaussian kernel $k(x, y) = \exp(-\lambda\, d_g(x, y)^2)$ is positive definite for all $\lambda > 0$, then M must have zero sectional curvature everywhere.*

Theorem 5 (Theorem 3 in Da Costa et al. (2023)) *On the circle S^1 with its geodesic distance, the Gaussian kernel is not positive definite for any $\lambda > 0$. More generally, this holds for any metric space into which S^1 can be isometrically embedded.*

The following is a folklore lemma that suggests that the Gaussian kernel may be positive definite on many spaces, as long as one chooses the bandwidth parameter λ to be large. Of course, Theorem 5 shows that it is not true that every space admits a choice of λ that makes it positive definite. The lemma holds because a Gram matrix for distinct points converges to the identity as $\lambda \to \infty$, and we omit the details here.

Lemma 6 *Let (X, d) be a metric space and $x_1, \dots, x_N \in X$ be a finite set of points. Then there exists $\lambda > 0$ such that the Gram matrix K for the points x_i and the Gaussian kernel with bandwidth parameter λ is positive semi-definite.*

3.3. Defining the Spaces Y_α

In this subsection we define the cone metric on a two-dimensional cone with cone angle α, denoted by (Y_α, d_c). These are important examples in metric geometry and Riemannian geometry, and a thorough introduction can be found in Burago et al. (2001), from which we take the definition of Y_α.

Definition 7 (Definition 3.6.16 in Burago et al. (2001)) *Let $P := \{(r, \theta) \in \mathbb{R}^2 : r \geq 0, \theta \in \mathbb{R}\}$. For $\alpha > 0$ define the equivalence relation $\sim$ via $(r, \theta) \sim (r, \theta + k\alpha)$ and $Y_\alpha := \mathbb{R}^2 / \sim$. Every point $P \in Y_\alpha$ has a representative (r, θ) with $\theta \in [0, \alpha)$. We then define a metric on Y_α via*

$$\begin{aligned} \overline{d}(a, b) &:= \min\{|a - b|, \alpha - |a - b|, \pi\} \text{ for } a, b \in \mathbb{R}, \\ d_c([(r_1, \theta_1)], [(r_2, \theta_2)])^2 &:= r_1^2 + r_2^2 - 2r_1 r_2 \cos(\overline{d}(\theta_1, \theta_2)), \text{ for } \theta_1, \theta_2 \in [0, \alpha). \end{aligned} \tag{3}$$

We conjecture that for some choices of α, the Gaussian kernel on the space Y_α is positive definite for some values of λ but not for all values of λ. Roughly speaking, that is because for $\alpha = 2\pi$ the space is isometric to Euclidean space, so the Gaussian kernel is positive definite for all values of λ, but for small α the space Y_α looks like a very thin cone that contains a circle (going around the cone) that is close to being isometric to the circle with the standard metric.

This is made precise in the next two lemmas:

Lemma 8 *On Y_α for $\alpha = 2\pi$, the Gaussian kernel is positive definite for all $\lambda > 0$.*

Proof For $\alpha = 2\pi$ we have that Equation 3 is the Euclidean distance on $\mathbb{R}^2$ for points in polar coordinates, i.e. $Y_{2\pi}$ is isometric to $\mathbb{R}^2$ and the claim follows from Theorem 3. ■

Lemma 9 *For every $\lambda > 0$ there exists $\alpha > 0$ such that the Gaussian kernel is not positive definite on Y_α.*

Proof Step 1: an approximate isometry from S^1 into Y_α.

Let $S^1 = [0, 1]/\sim$ be the unit interval with end points identified. Define $F : S^1 \to Y_\alpha$ as $F(t) = (1/\alpha, \alpha t)$. Assume $\alpha < \pi$. Then F is close to being an isometry in the following sense. For $s, t \in S^1$:

$$\begin{aligned} d_c(F(s), F(t))^2 = \frac{2}{\alpha^2}(1 - \cos(\alpha d_{S^1}(s, t)) &= \frac{2}{\alpha^2}\left(\frac{(\alpha \cdot d_{S^1})^2}{2} - \frac{(\alpha \cdot d_{S^1})^4}{24} + O(\alpha^6)\right) \\ &= d_{S^1}(t_i, t_j)^2 - \frac{\alpha^2}{12} d_{S^1}(t_i, t_j)^4 + O(\alpha^4), \end{aligned} \tag{4}$$

where we used the Taylor expansion $1 - \cos(x) = \frac{x^2}{2} - \frac{x^4}{24} + O(x^6)$.

Step 2: using the result from S^1.

Let $\lambda > 0$. By Theorem 5 there exist points $x_1, \dots, x_N$ so that the Gram matrix K has a negative eigenvalue. Denote the Gram matrix of $F(x_1), \dots, F(x_N)$ by $K_F(\alpha)$. By Equation 4 we have that $K_F(\alpha) \to K$ as $\alpha \to 0$. Thus, for α small enough, there exists a negative eigenvalue of $K_F(\alpha)$, which proves the claim. ■

Taken together, these two results suggest the following conjecture:

Conjecture 10 *For $\alpha \in (0, 2\pi)$, the Gaussian kernel on (Y_α, d_c) is positive definite if λ is large compared to α^{-1}.*

4. Experiments

To test the relationship between the cone angle α and the parameter λ, we numerically test for positive definiteness. We randomly sample points in Y_α and compute the eigenvalues of the resulting Gram matrix of the Gaussian kernel for various choices of α and λ.

Table 1 and Figure 1 summarise the results: the table shows if there were points found for which the Gram matrix has a negative eigenvalue. The figure shows for how many out of the 100 iterations a negative eigenvalue was found. We remark that even if no negative eigenvalue was found, this is no proof that the Gaussian kernel is positive definite.

Our experiments confirm the theoretical results and conjecture from above, which we sum up in the following:

1. For $\alpha = 2\pi$, our experiments agree with Theorem 3, confirming that the Gaussian kernel is positive definite for all values of λ.

2. For $\alpha < 2\pi$ we find that the Gaussian kernel is not positive definite for fixed λ and small α (which is suggested by theorem 9), but it may be positive definite for small α, if one takes λ sufficiently large, cf. Conjecture 10. For small α, the curvature of Y_α is large in the sense of Alexandrov spaces (see (Burago et al., 2001, Example 4.1.4)), suggesting a quantitative relation between the range of positive bandwidth parameters and curvature.

3. For $\alpha > 2\pi$ we know by theorem 4 that the Gaussian kernel is not positive for all values of λ, but our experiments suggest that it is positive for some values of λ. Our data neither suggest nor rule out a relation between the range of positive bandwidths parameters and α.

5. Conclusion

It is an open problem to understand under which conditions on the geometry the Gaussian kernel on metric spaces is positive definite. We studied the metric cones Y_α and our experiments suggest that the Gaussian kernel is positive definite on them for some values of λ but not all of them. Our experiments are compatible with the theoretical results in Schoenberg (1938b); Da Costa et al. (2023); Feragen et al. (2015), validating our numerics. They suggest a relation between the range of positive bandwidth parameters and the curvature of the underlying space. Studying classes of model spaces with precisely understood geometry, such as our Y_α, may be a route to a complete understanding of how positivity of the Gaussian kernel is related to curvature and topology of the underlying space.

References

N. Aronszajn. Theory of reproducing kernels. *Transactions of the American Mathematical Society*, 68(3):337–404, 1950.

Arit Kumar Bishwas, Ashish Mani, and Vasile Palade. Gaussian kernel in quantum learning. *International Journal of Quantum Information*, 18(03):2050006, 2020.

Dmitri Burago, Yuri Burago, and Sergei Ivanov. *A course in metric geometry*, volume 33 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001.

Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine Learning*, 20(3): 273–297, 1995.

Nathaël Da Costa, Cyrus Mostajeran, and Juan-Pablo Ortega. The gaussian kernel on the circle and spaces that admit isometric embeddings of the circle. In *International Conference on Geometric Science of Information*, pages 426–435. Springer, 2023.

Yutao Feng, Xiang Feng, Yintong Shang, Ying Jiang, Chang Yu, Zeshun Zong, Tianjia Shao, Hongzhi Wu, Kun Zhou, Chenfanfu Jiang, et al. Gaussian splashing: Unified particles for versatile motion synthesis and rendering. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 518–529, 2025.

Aasa Feragen, Francois Lauze, and Søren Hauberg. Geodesic exponential kernels: When curvature and linearity conflict. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.

Siran Li. Gaussian kernels on nonsimply connected closed Riemannian manifolds are never positive definite. *Bulletin of the London Mathematical Society*, 56(1):263–273, 2024.

Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006.

I. J. Schoenberg. Metric spaces and completely monotone functions. *Annals of Mathematics*, 39(4):811–841, 1938a.

Isaac J Schoenberg. Metric spaces and positive definite functions. *Transactions of the American Mathematical Society*, 44(3):522–536, 1938b.

Bernhard Schölkopf and Alexander J. Smola. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. MIT Press, 2002.

B. W. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, 1986.

Gaochao Song, Chong Cheng, and Hao Wang. Gvkf: Gaussian voxel kernel functions for highly efficient surface reconstruction in open scenes. *Advances in Neural Information Processing Systems*, 37:104792–104815, 2024.

Suvrit Sra. Positive definite matrices and the s-divergence. *Proceedings of the American Mathematical Society*, 144(7):2787–2797, 2016.

Shasha Yuan, Weidong Zhou, Qi Wu, and Yanli Zhang. Epileptic seizure detection with log-euclidean gaussian kernel-based sparse representation. *International journal of neural systems*, 26(03):1650011, 2016.

Appendix A. Experimental results

In this section we present the numerical results of the experiments explained in section 4. For each combination of α and λ we randomly sampled 1500 points and checked the corresponding Gram matrix and repeated this process 100 times. Rounding errors can yield negative eigenvalues close to zero, even if the Gram matrix is positive semidefinite. Because of this, we checked for negative eigenvalues smaller than -10^{-4}.

If a negative eigenvalue is found, this *proves* that the Gaussian kernel on this space for this bandwidth parameter is *not positive definite*. If, after 100 iterations, no negative eigenvalue is found, this is *no proof* that the Gaussian kernel is positive definite, but it suggests that may be the case.

In Table 1, we show all tested combinations of α and λ and whether a negative eigenvalue of the Gram matrix was found (red colour) or no negative eigenvalue was found (white colour).

$\alpha \setminus \lambda$	**0.1**	**0.6**	**1**	**5.6**	**10**	**20**	**35**	**50**	**80**	**100**	**250**	**300**	**450**	**500**	**700**	**800**	**1000**	**1500**	**1700**	**2500**	**5000**
0.1π																					
0.3π																					
0.5π																					
0.7π																					
0.9π																					
π																					
1.1π																					
1.3π																					
1.5π																					
1.7π																					
1.9π																					
2π																					
2.5π																					
3π																					
3.6π																					
4.7π																					
5π																					
7π																					
7.8π																					
8.8π																					
10π																					

Table 1: experimental results of Gaussian kernel positive definiteness for different values of α and λ. A red cell indicates that a Gram matrix with negative eigenvalue was found, thereby *proving* that the Gaussian kernel is not positive definite for this choice of α and λ. A white cell indicates that after 100 iterations no negative eigenvalue was found. This suggests that for this choice of α and λ the Gaussian kernel *may* be positive definite, but it is no proof of that fact.

In Figure 1 we show the same combinations of α and λ and display for what percentage of the 100 iterations we found a negative eigenvalue in the Gram matrix. If at least in one iteration a negative eigenvalue was found, then the corresponding cell in Table 1 is red. The figure suggests that if, for a choice of α and λ, the Gaussian kernel is close to being positive definite, then the probability for randomly chosen points to produce a Gram matrix with negative eigenvalue is small.

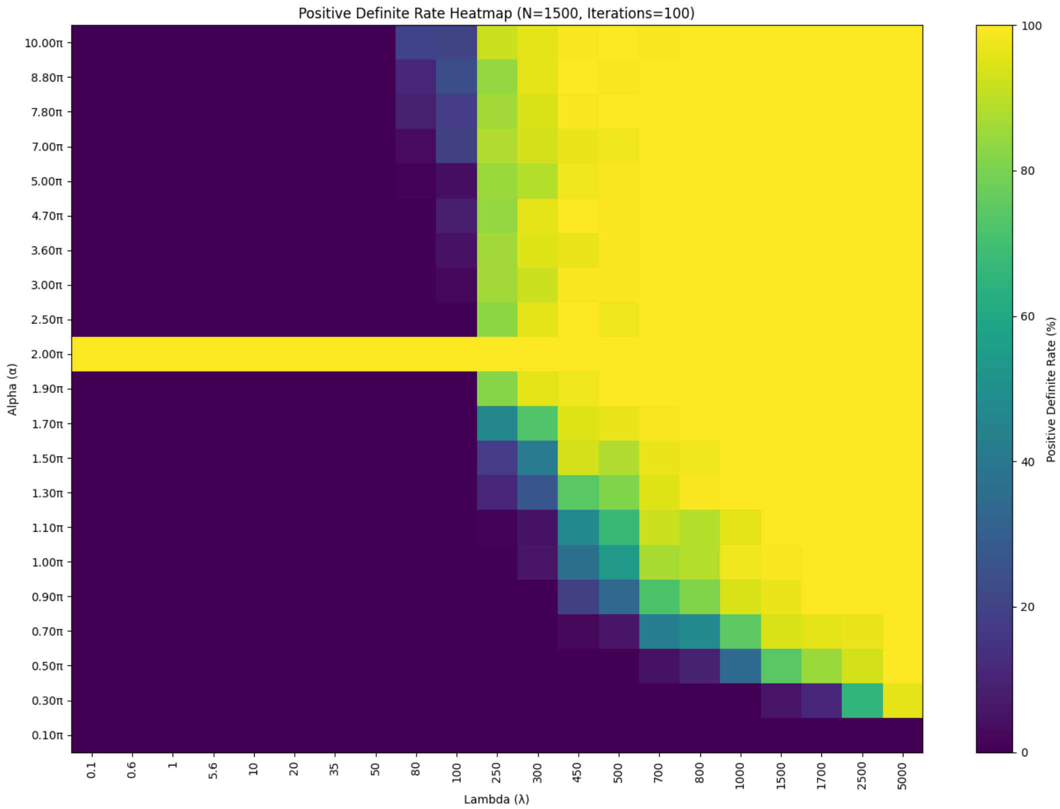

Figure 1: heatmap showing which percentage of iterations yielded a negative eigenvalue in the Gram matrix constructed from 1500 randomly sampled points. The data in this heatmap is a refinement of the data presented in Table 1, in the sense that a cell in Table 1 is white if and only if it is completely yellow in this heatmap, i.e. zero out of 100 iterations produced a negative eigenvalue in the Gram matrix.

To provide a more detailed view beyond the summary statistics in Table 1 and the heatmap in Figure 1, Figure 2 visualizes the full eigenspectrum of the Gram matrix for four representative scenarios. Each plot displays all 1500 eigenvalues, sorted in descending order, for a specific choice of cone angle α and kernel parameter λ. An inset in each panel provides a magnified view of the smallest eigenvalues, allowing for a clear visual inspection of the kernel's positive definiteness. This is the same type of visualisation as in (Feragen et al., 2015, Figure 5).

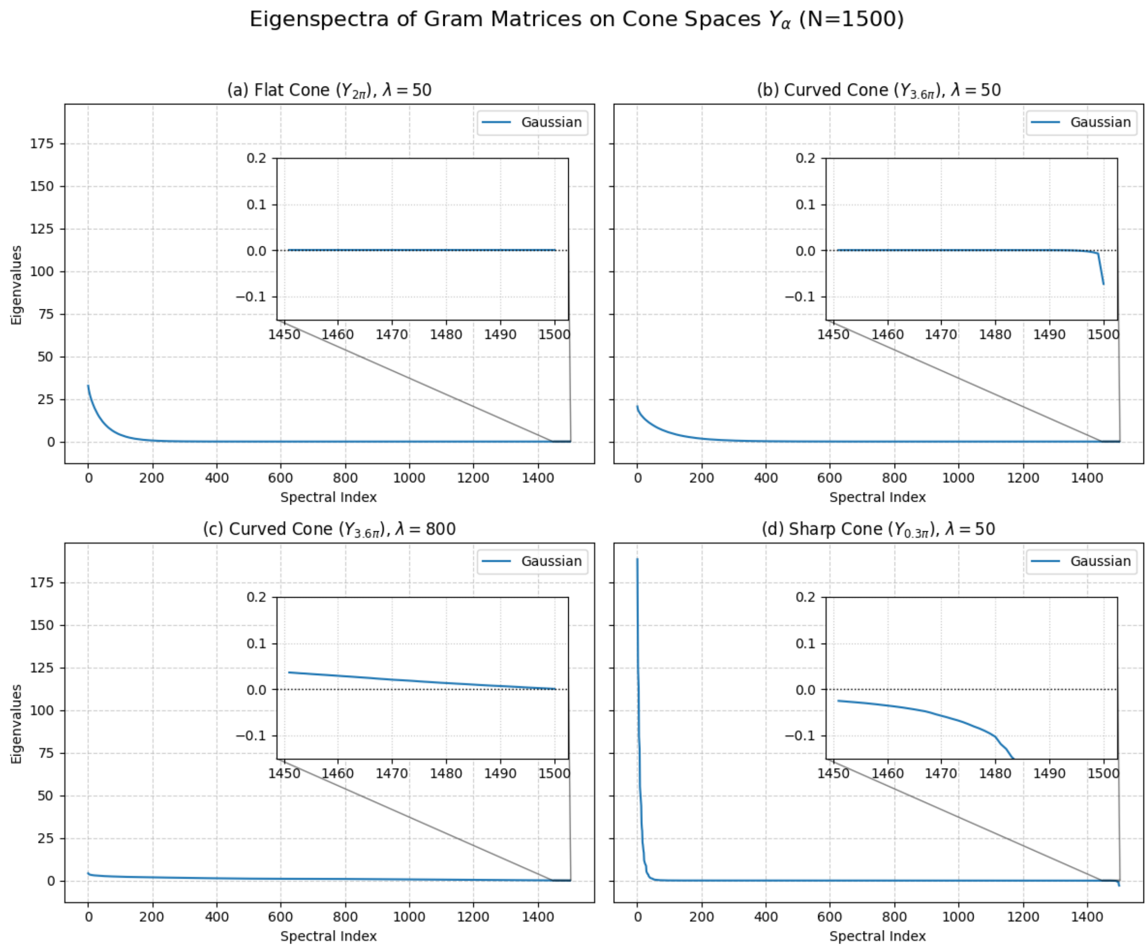

Figure 2: the plots show the eigenspectra of Gram matrices for the Gaussian kernel on spaces Y_α for different choice of α and λ and for 1500 randomly chosen points. (a): The kernel is known to be positive definite on $Y_{2\pi}$ by theorem 3. (b): For $Y_{3.6\pi}$ the kernel is known to not be positive definite for all λ by Feragen et al. (2015) and this plot suggests this is the case for $\lambda = 50$. (c): It is not known whether there are values of λ for which the kernel is positive definite, but our experiments suggest this may be the case for $\lambda = 800$. (d): Theorem 9 suggests that the kernel for small α the Gaussian kernel is not positive definite, even if λ is large.

Appendix B. A discrete metric space on which the Gaussian kernel is positive definite for some values of λ

It was remarked in (Da Costa et al., 2023, Section 3) and (Li, 2024, Section 3) that one can construct discrete metric spaces on which the Gaussian kernel is positive definite for some values of λ but not for all of them. In this section we write down one such example.

Let $X = \{x_1, x_2, x_3, x_4\}$ be a metric space with distances given by $d(x_1, x_i) = 1$ for $i = 2, 3, 4$, and $d(x_i, x_j) = 2$ for distinct $i, j \in \{2, 3, 4\}$.

Lemma 11 *The Gaussian kernel of (X, d) is positive definite for $\lambda \geq \frac{\ln(2)}{2}$ and not positive definite otherwise.*

Proof When checking Gram matrices, it suffices without loss of generality to check Gram matrices constructed from pairwise distinct points. The associated Gaussian kernel matrix $K_{ij} = \exp(-\lambda d(x_i, x_j)^2)$ is

$$K = \begin{pmatrix} 1 & x & x & x \\ x & 1 & y & y \\ x & y & 1 & y \\ x & y & y & 1 \end{pmatrix},$$

where $x = e^{-\lambda}$ and $y = e^{-4\lambda} = x^4$. For the kernel to be positive definite, all its eigenvalues must be positive.

The eigenvalues of K are $1 - y$ (with multiplicity 2) and $1 + x^4 \pm x\sqrt{x^6 + 3}$. Since $\lambda > 0$, we have $x \in (0, 1)$, which ensures that the first two eigenvalues, $1 - y = 1 - x^4$, are positive. The third eigenvalue, $1 + x^4 + x\sqrt{x^6 + 3}$, is also clearly positive. Positive definiteness therefore hinges on the sign of the fourth eigenvalue,

$$\mu = 1 + x^4 - x\sqrt{x^6 + 3}.$$

The condition $\mu \geq 0$ is equivalent to $1 + x^4 \geq x\sqrt{x^6 + 3}$. As both sides are positive, we can square them to obtain $2x^4 - 3x^2 + 1 \geq 0$, which factors as $(2x^2 - 1)(x^2 - 1) \geq 0$. Since $x \in (0, 1)$, the term $x^2 - 1$ is negative. The inequality thus holds if and only if $2x^2 - 1 \leq 0$, which is equivalent to $e^{-2\lambda} \leq 1/2$. This simplifies to

$$\lambda \geq \frac{\ln(2)}{2}.$$

■