# Graph and reproducing kernel Hilbert spaces

Mohammadreza Foroutan[1]
Department of Mathematics, Payame Noor University, P.O.Box 19395-3697, Tehran, Iran

Farzad Farzanfar
Department of Computer Engineering and Information Technology, Payame Noor University,
P.O. Box 19395-3697 Tehran, Iran

**Abstract**

In this paper, We study reproducing kernels, and associated reproducing kernel Hilbert spaces over infinite, discrete and countable sets V. By using composition operators defined an injective homomorphisms on infinite weighted graphs from a viewpoint of reproducing kernel Hilbert space theory. Also, we study the main notions and tools we shall need for our graph analysis; this includes the theory of weighted networks, reproducing kernel and graph Laplacians.

**Keywords:** Reproducing kernels, Graph Laplacians, Injective homomorphisms, Infinite weighted graphs

**AMS Mathematical Subject Classification [2010]:** 05C50, 46E22

## 1 Introduction

The concept of a graph is one of the most fundamental mathematical concepts ever conceived. Graph has found many applications in engineering and science, such as chemical, civil, electrical and mechanical engineering, architecture, management and control, communication, operational research, sparse matrix technology, combinatorial optimisation, and computer science [7, 10]. In mathematics, graphs are unavoidable as they appear (implicitly) whenever there is a relation between objects. In particular, they play a most prominent role in various combinatorial questions. At the same time, graphs often come about via approximation schemes when dealing with a continuous setting.

Within computer science, cybernetics uses graphs to represent networks of communication, data organization, computational devices, the flow of computation, etc. For instance, the link structure of a website can be represented by a directed graph, in which the vertices represent web pages and directed edges represent links from one page to another. A similar approach can be taken to problems in social media,[6] travel, biology, computer chip design, mapping the progression of neuro-degenerative diseases,[12, 13] and many other fields. The development of algorithms to handle graphs is therefore of major interest in computer science. The transformation of graphs is often formalized and represented by graph rewrite systems. Complementary to graph transformation systems focusing on rule-based in-memory manipulation of graphs are

[1]speaker

graph databases geared towards transaction-safe, persistent storing and querying of graph-structured data. In this paper, we focus on two important topics in graph-structured data analysis: graph comparison and graph matching, for all of which we propose effective algorithms by making use of kernel functions and the corresponding reproducing kernel Hilbert spaces [8, 10]. Graph kernels, which are positive definite functions on graphs, are powerful similarity measures, in the sense that they make various kernel-based learning algorithms, for example, clustering, classification, and regression, applicable to structured data. Our graph kernels are obtained by two-step embeddings. In the first step, we represent the graph nodes with numerical vectors in Euclidean spaces. In the second step, we represent the whole graph with an element in reproducing kernel Hilbert spaces. The experimental results show that our graph kernels significantly outperform state-of-the-art approaches in both accuracy and computational efficiency.

## 2 Reproducing Kernel Hilbert Spaces

A reproducing kernel Hilbert space is a Hilbert space $H$ of functions on a prescribed set, say $X$ with the property that point-evaluation for functions $f \in H$ is continuous with respect to the $H$-norm. They are called kernel spaces, because, for every $x \in X$, the point-evaluation for functions $f \in H$, $f(x)$ must then be given as a $H$-inner product of $f$ and a vector $k_x$, in $H$; called the kernel [5, 4].

**Definition 2.1.** [3] Let $X$ be a set, and $\mathcal{F}(X)$ denotes the set of all finite subset of $X$. A function $k : X \times X \longrightarrow \mathbb{C}$ is said to be positive definite, if

$$\sum\sum\nolimits_{(x,y)\in F\times F}\overline{c_x}c_y k(x,y) \geq 0,$$

holds for all coefficients $\{c_x\}_{x\in F} \subset \mathbb{C}$, and all $F \in \mathcal{F}(X)$.

**Definition 2.2.** Fix a countable infinite set $X$,

(1) For all $x \in X$, set $k_x := k(\cdot, x) : X \to \mathbb{C}$ as a function on $X$.

(2) Let $H := H(k)$ be the Hilbert-Completion of the *span*$\{k_x : \ x \in X\}$, with respect to the inner product

$$\langle \sum\nolimits_{x\in F} c_x k_x, \sum\nolimits_{y\in F} d_y k_y\rangle_H := \sum\sum\nolimits_{(x,y)\in F\times F}\overline{c_x}d_y k(x,y), \quad F \in \mathcal{F}(X),$$

modulo the subspace of functions of zero $H$-norm, i.e,

$$\sum\sum\nolimits_{(x,y)\in F\times F} k(x,y)\overline{c_x}c_y = 0.$$

$H$ is then a reproducing kernel Hilbert space, with the reproducing property:

$$\langle k_x, f\rangle_H = f(x) \quad \forall x \in X,\ \forall f \in H.$$

**Theorem 2.3.** *$H$ is a reproducing kernel Hilbert space (i.e., its evaluation operator $\delta_x$ are bounded linear operators), if and only if $H$ has a reproducing kernel.*

*Proof.* We show that, if $H$ has a reproducing kernel then $\delta_x$ is a bounded linear operators,

$$\begin{aligned}|\delta_x[f]| = |f(x)| = |\langle f, k(.,x)\rangle_H| &\leq \|k(.,x)\|_H\|f\|_H \\ &= \langle k(.,x), k(.,x)\rangle_H^{\frac{1}{2}}\|f\|_H = k^{\frac{1}{2}}(x,x)\|f\|_H,\end{aligned}$$

where Cauchy- Schwarz used in 3rd line. Consequently $\delta_x : \mathcal{F} \to \mathbb{R}$ bounded with $\lambda_x = k^{\frac{1}{2}}(x, x)$.
Now, we show that, if $\delta_x$ be a bounded linear operators, then $H$ has a reproducing kernel. Suppose $\delta_x$ be a bounded linear operators, then by Riesz representation exists $f_{\delta_x} \in H$ such that

$$\delta_x(f) = \langle f, f_{\delta_x} \rangle_H, \quad \forall\, f \in H.$$

Define $k(., x) = f_{\delta_x}(.)$, $\forall\, x \in X$. By definition of $k(., x)$ and $f_{\delta_x}$, we have $k(., x) = f_{\delta_x}(.) \in H$ and $\langle f(.), k(., x) \rangle_H = \delta_x(f) = f(x)$. Therefore, $k$ is the reproducing kernel. □

**Theorem 2.4.** *$k$ is a positive definite kernel on the set $X$ if and only if there exists a Hilbert space $H$ and a mapping $\phi : X \to H$, such that, for any $x, x^{'}$ in $X$:*

$$k(x, x^{'}) = \langle \phi(x), \phi(x^{'}) \rangle_H.$$

*Proof.* See proof in [2]. □

**Remark 2.5.** Let $d_k(x_1, x_2) = \|\phi(x_1) - \phi(x_2)\|_H$, then

$$\begin{aligned} d_k^2(x_1, x_2) &= \|\phi(x_1) - \phi(x_2)\|_H^2 \\ &= \langle \phi(x_1) - \phi(x_2), \phi(x_1) - \phi(x_2) \rangle_H \\ &= \langle \phi(x_1), \phi(x_1) \rangle_H + \langle \phi(x_2), \phi(x_2) \rangle_H - 2\langle \phi(x_1), \phi(x_2) \rangle_H \\ &= k(x_1, x_1) + k(x_2, x_2) - 2k(x_1, x_2). \end{aligned}$$

**Remark 2.6.** $-d_k^2$ is conditionally positive definite, when for all $t > 0$, $exp(-td_k^2(x, x^{'}))$ is positive definite.

**Notations**:

- A directed graph is a pair $G = (V, E)$ with $V$ finite vertices and $E \subset V \times V$ (edges).
- A graph is labeled if a label from a set of labels $A$ is assigned to each vertex and/or edge.
- Two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ are isomorphic (denoted $G_1 \cong G_2$) if there exists a bijection between $V_1$ and $V_2$ that preserves edges and labels.

**Definition 2.7.** We note $\mathcal{G}$ the quotient set of the set of all labelled graphs with respect to isomorphism.
A graph kernel is a positive definite kernel over $\mathcal{G}$.
A graph kernel is complete if it separates non-isomorphic graphs,i.e.,

$$\forall\, G_1, G_2 \in \mathcal{G},\ \ d_k(G_1, G_2) = 0 \ \Rightarrow G_1 \cong G_2,$$

where $d_k^2(G_1, G_2) = k(G_1, G_1) + k(G_2, G_2) - 2k(G_1, G_2)$. Equivalently, $\phi(G_1) \neq \phi(G_2)$ if $G_1$ and $G_2$ are not isomorphic.

Given a graph $G = (V, E)$ the graph distance $d_G(x, x^{'})$ between any two vertices is the length of the shorts path between $x$ and $x^{'}$.
We say that the graph $G = (V, E)$ can be embedded in a Hilbert space if $-d_G$ is conditionally positive definite, which implies in particular that $exp(-td_G(x, x^{'}))$ is positive definite for all $t > 0$.

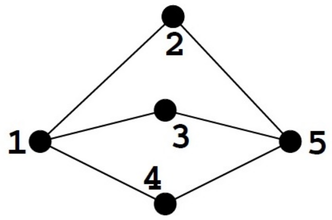


Figure 1: Graph of Example 2.8

**Example 2.8.** Following graph is not conditionally positive definite graph distance.
According to the Figure 1 we have

$$d_G = \begin{pmatrix} 0 & 1 & 1 & 1 & 2 \\ 1 & 0 & 2 & 2 & 1 \\ 1 & 2 & 0 & 2 & 1 \\ 1 & 2 & 2 & 0 & 1 \\ 2 & 1 & 1 & 1 & 0 \end{pmatrix},$$

$$\lambda_{\min}\big(exp(-0.2d_G(i,j)\big) = -0.028 < 0.$$

The adjacency matrix of graph $G = (V, E)$ is a square $A$ whose rows and columns are indexed by $V$ with $(x, y)$-entry 1 if $\{x, y\} \in E$ and 0 otherwise. The degree matrix $D$ of $G$ is a diagonal matrix with $(x, x)$-entry equal to its degree. The Laplacian matrix $L$ of $G$ is defined to be $L = D - A$. The set $H = \{f \in \mathbb{R}^n : \sum_{i=1}^n f_i = 0\}$ endowed with norm

$$\|f\| = \sum_{i\sim j}(f(x_i) - f(x_j))^2,$$

is a reproducing kernel Hilbert space with reproducing kernel $L^*$, where $L^*$ is a pseudo-inverse of the graph Laplacian. For example, in the following graph, we have

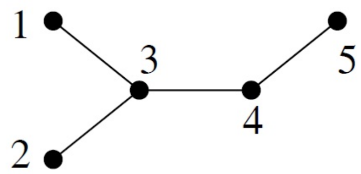


Figure 2:

$$A = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \qquad D = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$L = A - D = \begin{pmatrix} 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ -1 & -1 & 3 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}, \qquad L^* = \begin{pmatrix} 0.88 & -0.12 & 0.08 & -0.32 & -0.52 \\ -0.12 & 0.88 & 0.08 & -0.32 & -0.52 \\ 0.08 & 0.08 & 0.028 & -0.12 & -0.32 \\ -0.32 & -0.32 & -0.12 & 0.48 & 0.28 \\ -0.52 & -0.52 & -0.32 & 0.28 & 1.08 \end{pmatrix}.$$

## 3 Reproducing kernel in weigh graphs

In this section, we study reproducing kernel in infinite graphs. Let $V$ be a infinite set. A graph over $V$ or a infinite graph is $G_{w,e}$ consisting of a function $w : V \times V \to [0, \infty)$ satisfying

- $w(x, y) = w(y, x)$, for all $x, y \in V$,
- $w(x, x) = 0$, for all $x \in V$,
- $\sum_{y \in V} w(x, y) < \infty$, for all $x \in V$,

and a function $e : V \to [0, \infty)$. If $e(x) = 0$ for all $x \in V$, then we speak of $G_w$ as a graph over $V$. The elements of $V$ are called the vertices of the graph. The map $w$ is called an edge weigh. More specifically, a pair $(x, y)$ with $w(x, y) > 0$ is called and edge with $w(x, y)$ connecting $x$ to $y$. The vertices $x$ and $y$ are called neighbors if they from an edge. In this case, we write $x \sim y$. The map $e$ is called the killing term.

**Example 3.1.** If $w$ take value in $\{0, 1\}$ and $e(x) = 0$ for all $x \in V$, then we speak of a graph $G_w$ with standard weights. In this case, the set of edge $E$ is given by

$$E = \{(x, y) \in V \times V \mid w(x, y) = 1\}.$$

An important geometric quantity that comes with graph $G_{w,e}$ over $V$ is the vertex degree. Let $G_{w,e}$ be a graph over a infinite set $V$, The degree is the function $deg : V \to [0, \infty)$ given by

$$deg(x) = \sum_{y \in V} w(x, y) + e(x).$$

Let $C(V)$ be the set of all real-valued functions on $V$ and $C_c(V)$ be the subset of all real-valued functions of finite support. To a graph $G_{w,e}$ over $V$ we associate the quadratic form

$$\Delta := \Delta_{w,e} : C(V) \to [0, \infty)$$

which acts by

$$\Delta(f) = \frac{1}{2} \sum_{x,y \in V} w(x, y)\big(f(x) - f(y)\big)^2 + \sum_{x \in V} f^2(x)e(x).$$

We will be interested in the space of all functions of finite energy, which is defined as

$$D := \{f \in C(V) \mid \Delta(f) < \infty \ \ and \ \ f(0_G) = 0\}.$$

Obviously the equality

$$\Delta(\delta_x) = deg(x),$$

holds. Therefore, assumption $\sum_{y \in V} w(x, y) < \infty$ implies that $C_c(V)$ is contained in $D$. Now. we extend $\Delta : D \times D \to \mathbb{R}$ to a bilinear map

$$\Delta(f, g) = \varepsilon(f, g) + \sum_{x \in V} f(x)g(x)e(x),$$

where $\varepsilon(f, g) = \frac{1}{2} \sum_{x,y \in V} w(x, y)\big(f(x) - f(y)\big)\big(g(x) - g(y)\big)$. Note that the above sum is absolutely convergent by the definition of $D$.

**Theorem 3.2.** *For any $x$ in $V$, there exists a unique function $k_x$ in $D$ such that $\langle f, k_x\rangle_D = f(x)$ for any $f$ in $D$, that is, $D$ is a reproducing kernel Hilbert space.*

*Proof.* We fix an arbitrary vertex $x$. Then there exists a finite path $P = \{x_0, x_1, \cdots, x_n\}$ from $x_0 = 0_G$ to $x_n = x$ in $G$ by assumption that $G$ is connected. Then we have that

$$\begin{aligned}|f(x)| &\leq \sum_{j=0}^{n-1}|f(x_j) - f(x_{j+1})| + \sum_{j=0}^{n}|f(x_j)|\\ &\leq \Big(\sum_{j=0}^{n-1}\frac{1}{w(x_j, x_{j+1})}\Big)^{\frac{1}{2}}\Big(\sum_{j=0}^{n-1}w(x_j, x_{j+1})|f(x_j) - f(x_{j+1})|^2\Big)^{\frac{1}{2}} + \Big(\sum_{j=0}^{n-1}\frac{1}{e(x_j)}\Big)^{\frac{1}{2}}\Big(\sum_{j=0}^{n-1}e(x_j)|f(x_j)|^2\Big)^{\frac{1}{2}}\\ &\leq c\|f\|_D\end{aligned}$$

by the Cauchy-Schwarz inequality. Hence evaluation at $x$ is bounded on $D$. By the Riesz representation theorem, we have the conclusion. □

Let $G_{w_1,e_1}$ and $G_{w_2,e_2}$ be graphs. A map $\phi$ from $V_1 = V(G_{w_1,e_1})$ into $V_2 = V(G_{w_2,e_2})$ is called a homomorphism of $G_{w_1,e_1}$ into $G_{w_2,e_2}$ if $w_1(x_1, y_1) \leq w_2(\phi(x_1), \phi(y_1))$ and $e_1(x) \leq e_2(\phi(x))$ for any $x_1, x_2, x \in V_1$. Furthermore, $G_{w_1,e_1}$ and $G_{w_2,e_2}$ are said to be isomorphic if there exists a bijective map $\phi$ between $V_1$ and $V_2$ with preserves adjacency, that is, both $\phi$ and $\phi^{-1}$ are homomorphisms.
Let $D_1$ and $D_2$ denote the reproducing kernel Hilbert spaces consisting of real-valued functions on $V_1$ and $V_2$ with norms $\|.\|_1$ and $\|.\|_2$ respectively, and $\phi$ be a homomorphism from graph $G_1$ into $G_2$. For each function $f \in H_2$, $S_\phi f = f \circ \phi$ defines a linear operator $S_\phi$ from $H_2$ into $H_1$.

**Theorem 3.3.** *Let $S_\phi : D_2 \to D_1$, $S_\phi f = f \circ \phi$, set $M_\phi = \max_{x_2 \in V_2}|\phi^{-1}(x_2)|$, then*

$$\|S_\phi f\|_{D_1} \leq M_\phi\|f\|_{D_2}.$$

*Proof.* For any $f \in D_2$, we have

$$\begin{aligned}\varepsilon_1(f\circ\phi, f\circ\phi) &= \frac{1}{2}\sum_{x_1,y_1\in V_1} w(x_1, y_1)|f\circ\phi(x_1) - f\circ\phi(y_1)|^2\\ &\leq \frac{1}{2}\sum_{x_1,y_1\in V_1} w\big(\phi(x_1), \phi(y_1)\big)|f\circ\phi(x_1) - f\circ\phi(y_1)|^2\\ &= \frac{1}{2}\sum_{x_2,y_2\in\phi(V_1)} w(x_2, y_2)|f(x_2) - f(y_2)|^2|\phi^{-1}(x_2)||\phi^{-1}(y_2)|\\ &\leq \frac{M_\phi^2}{2}\sum_{x_2,y_2\in V_2} w(x_2, y_2)|f(x_2) - f(y_2)|^2\\ &= \frac{M_\phi^2}{2}\varepsilon_2(f, f),\end{aligned}$$

and

$$\begin{aligned}\sum_{x_1\in V_1}|f\circ\phi(x_1)|^2 e(x_1) &\leq \sum_{x_1\in V_1}|f\circ\phi(x_1)|^2 e\big(\phi(x_1)\big)\\ &= \sum_{x_2\in\phi(V_1)}|f(x_2)|^2|\phi^{-1}(x_2)|e(x_2)\\ &= M_\phi\sum_{x_2\in V_2}|f(x_2)|^2 e(x_2).\end{aligned}$$

These inequalities conclude that $\|S_\phi f\|_{D_1} \leq M_\phi\|f\|_{D_2}$. □

## References

[1] N. Aronszajn, *Reproducing and pseudo-reproducing kernels and their application to the partial differential equations of physics. Studies in partial differential equations*, Technical report 5, preliminary note. Harvard University, Graduate School of Engineering., 1948.

[2] M. Cui and Y, Lin, *Nonlinear Numerical Analysis in the Reproducing Kernel Space*, Nova Science New York, 2009.

[3] M. R. Foroutan, A. Ebadian and R. Asadi, *Reproducing kernel method in Hilbert spaces for solving the linear and nonlinear four-point boundary value problems* , Int. J. Comput. Math., 95(10) (2018), 2128–2142.

[4] M. R. Foroutan, A. S. Gholizadeh, Sh. Najafzadeh, R. H. Haghi, *Laguerre reproducing kernel method in Hilbert spaces for unsteady stagnation point flow over a stretching/shrinking sheet*, Appl. Math. J. Chinese Univ., 36(3) (2021), 354–-369.

[5] M. R. Foroutan, M. S. Hashemi, L. Gholizadeh, A. Akgül, F. Jarad, *A new application of the Legendre reproducing kernel method* , AIMS Mathematics, 7(6) (2022), 10651–-10670.

[6] M. Grandjean, *A social network analysis of Twitter: Mapping the digital humanities community*, Cogent Arts and Humanities, 3(1) (2016), 1171458.

[7] P. E. T. Jorgensen, M. S. Song, *Compactification of infinite graphs and sampling*, Theory Signal Image Process., 12(2-3) (2013), 139–-158.

[8] P. E. T. Jorgensen, F. Tian, *Discrete reproducing kernel Hilbert spaces: sampling and distribution of Dirac-masses*, J. Mach. Learn. Res., 16 (2015), 3079-–314.

[9] P. E. T. Jorgensen, F. Tian, *Positive definite kernels and boundary spaces*, Adv. Oper. Theory, 1 (2016), 123-–133.

[10] P. E. T. Jorgensen, F. Tian, *Graph Laplacians and discrete reproducing kernel Hilbert spaces from restrictions*, Stochastic Analysis and Applications 34, 4 (2016), 722–747.

[11] P. Soardi, *Potential theory on infinite networks, Lecture Notes in Mathematics*, 1590. Springer-Verlag, Berlin, 1994.

[12] F. Vecchio, *Small World" architecture in brain connectivity and hippocampal volume in Alzheimer's disease: a study via graph theory from EEG data*, Cogent Arts and Humanities, 11(2) (2017), 473–485.

[13] F. Vecchio, *Brain network connectivity assessed using graph theory in frontotemporal dementia*, Neurology, 81(2) (2013), 134-–143.

e-mail: *mr_forootan@pnu.ac.ir*, *foroutan_mohammadreza@yahoo.com*
e-mail: *farzanfar@pnu.ac.ir*