Can the topology of two graphs be compared by one number?

P. Van Mieghem*

Delft University of Technology
5 June 2013 = 7^2

Abstract

Invoking the concept of a signature of a (sub)graph, we investigate how two (sub)graphs topologically can be compared by one number. We argue that one meaningful number cannot be found, unless the node labels in the graphs are specified and the same for the graphs.
It would be great to find the minimum number of bits that represents a graph with N nodes exactly and uniquely.

1 Signature of a graph

A graph $G(N, L)$ with N nodes and L links is usually specified [1] by an $N \times N$ adjacency matrix A with a zero diagonal and whose elements are either zero or one. Any $N \times N$ matrix A can be stored (and considered) as an $N^2 \times 1$ vector $vec(A)$, by concatenation of the rows (or columns) of the matrix (see [1, p. 254]). Hence, the topology of each (unweighted) graph is described by $N(N-1)$ bits. However, the vector of those $N(N-1)$ bits is not unique, because each permutation of the N node labels results in a different adjacency matrix, while the graph is still the same. Since the eigenvalues of a matrix are invariant to a similarity transform (such as a relabeling of nodes), the spectrum (set of N eigenvalues of A) is often used to discriminate between graphs. When confining to undirected graphs (for which A is symmetric and the eigenvalues are real), we need now N real numbers (or αN bits, where α is the number of bits of a real number in a computer), in contrast to $\frac{1}{2}N(N-1)$ bits. Hence, only if $N > 2\alpha + 1$, the spectrum is a more economical *discriminator* than the adjacency vector.

We define the *signature* of a graph $G(N, L)$ as a $r \times 1$ real vector σ that represents the graph uniquely (with overwhelming probability[1]). Before proceeding, we emphasize that a signature does not necessarily contain enough information to reproduce the graph G. Rather, the signature maps the graph space $\mathcal{G}$ containing all graphs to a signature space $\mathcal{S}$ of lower dimension so that information is lost to construct a unique inverse map from $\mathcal{S} \to \mathcal{G}$. To some extent, the signature can be compared to a hash function.

*Faculty of Electrical Engineering, Mathematics and Computer Science, P.O Box 5031, 2600 GA Delft, The Netherlands; *email*: P.F.A.VanMieghem@tudelft.nl

[1] It seems to be a difficult (open?) problem to find the minimum number of bits that represents a graph with N nodes exactly and uniquely.

Signatures may be used to compare properties among different graphs. For example, the influence of topology on the function of a biological network is discussed in [2]. Another question often arises: "Is a particular subgraph representative for the graph?"

1.1 Eigenvalue signature

Although cospectral graphs[2] exist for which Godsil and McKey have proposed an ingenious construction [3], Van Dam and Haemers [4] claim that the spectrum of the graph is a unique fingerprint when the size N of the graph is sufficiently large (larger than $N = 20$). We confine the discussion to the spectrum (eigenvalues only, without corresponding eigenvectors) of the adjacency matrix A. Inspired by the claim of van Dam and Haemers [4], a first proposal of a signature is the eigenvalue signature

$$\sigma_\lambda = (\lambda_N, \lambda_{N-1}, \ldots, \lambda_1) \tag{1}$$

where λ_k is the k-th largest eigenvalue of the adjacency matrix A and $r = N$ (but effectively $N - 1$ because $\sum_{j=1}^{N} \lambda_j = 0$ so that always one eigenvalue can be eliminated). Instead of the adjacency matrix, other graph theoretic matrices (such as the Laplacian Q) can be considered. In principle, any reordering of the N eigenvalues can be regarded as a valid signature. In order to reconstruct the graph or adjacency matrix A, we also need all eigenvectors x_k belonging to eigenvalue λ_k since $A = X\mathrm{diag}(\lambda_k)\, X^T$, where X is the matrix containing as columns the eigenvectors. Thus, the eigenvalue signature σ_λ does not contain enough information to reconstruct the graph.

We can construct *infinitely* many signatures based on σ_λ, that all are equally unique. Indeed, for any monotonous increasing function f, a generalized eigenvalue signature is

$$\sigma_{f(\lambda)} = (f(\lambda_N), f(\lambda_{N-1}), \ldots, f(\lambda_1))$$

which corresponds to $f(A) = X\mathrm{diag}(f(\lambda_k))\, X^T$, where the matrix X with eigenvectors is the same for any function f. Each generalized signature $\sigma_{f(\lambda)}$ still contains $r = N$ vector components; thus $\dim \sigma_{f(\lambda)} = N$.

The eigenvalue signature can be applied to weighted graphs with adjacency matrix W, where w_{ij} is a real (mostly non-negative) number, whereas a_{ij} is either 0 or 1. The uniqueness of σ_{λ_W} for a weighted adjacency matrix W is expected to hold more generally than for A as argued in [1, p. 249-250], because the number of cospectral weighted graphs is relatively smaller than the number of cospectral unweighted graphs.

1.2 Characteristic polynomial signature

A second proposal for a signature is the polynomial signature, in which each component is a function evaluation of the characteristic polynomial [1, p. 211-212]

$$c_A(x) = \prod_{j=1}^{N} (\lambda_j - x) = \det(A - xI) \tag{2}$$

[2] Cospectral graphs are non-isomorphic graphs with the same adjacency eigenvalues. Isomorphic graphs describe actually the same graph, only with a different node labeling (nodal permutation).

so that

$$\sigma_p = \left(c_A\left(x_1\right), c_A\left(x_2\right), \ldots, c_A\left(x_r\right)\right) \tag{3}$$

Each argument x_j with $1 \leq j \leq r$ can, in principle, be chosen arbitrarily, as well as the value of r. However, in order to show that σ_p is also a unique fingerprint for G, we must choose $r \geq N+1$ and require that all arguments x_j are different. These requirements are based on Lagrange's interpolation theorem [1, p. 274-275]. Indeed, the Lagrange interpolation polynomial that passes through the r points $\left\{\left(x_j, c_A\left(x_j\right)\right)\right\}_{1\leq j\leq r}$ is unique and of degree $r-1$ and to reconstruct all eigenvalues $\lambda_N, \lambda_{N-1}, \ldots, \lambda_1$, which are the zeros of the characteristic polynomial $c_A\left(x\right)$, the degree of the Lagrange polynomial should be at least equal to $r = N+1$. When $r > N+1$, then the Lagrange polynomial that passes through the r points $\left\{\left(x_j, c_A\left(x_j\right)\right)\right\}_{1\leq j\leq r}$ is

$$p_{r-1}\left(x\right) = \sum_{j=1}^{r} c_A\left(x_j\right) \frac{F_r\left(x\right)}{\left(x-x_j\right) F_r'\left(x_j\right)}$$

where $F_r\left(x\right) = \prod_{j=1}^{r}\left(x-x_j\right)$. The degree of the Lagrange polynomial $p_{r-1}\left(x\right)$ seems to be equal to $r-1 > N$. However, the difference polynomial

$$q_{r-1}\left(x\right) = p_{r-1}\left(x\right) - c_A\left(x\right)$$

is at most of degree $r-1$ and possesses r zeros at $x = x_j$ for $1 \leq j \leq r$, which is impossible because any polynomial of degree $r-1$ has precisely $r-1$ complex zeros. Thus, $q_{r-1}\left(x\right)$ must be zero and the Lagrange polynomial equals $p_{r-1}\left(x\right) = c_A\left(x\right)$ with degree $r = N+1$.

Using more sampling points than $N+1$ to determine a polynomial $c_A\left(x\right)$ of degree N cannot lead to a Lagrange polynomial of a higher degree than N. Also, using more information than necessary does not degrade the Lagrange polynomial in the sense that $p_{r-1}\left(x\right)$ is still precisely equal to $c_A\left(x\right)$. This property of Lagrangian sampling is particularly useful when we possess a set of r points $\left\{\left(x_j, f\left(x_j\right)\right)\right\}_{1\leq j\leq r}$ of which it is unknown whether $f\left(x\right)$ is a polynomial. Indeed, if $f\left(x\right)$ is a polynomial of degree N, then after generating more than $r > N+1$ function evaluations, the Lagrange polynomial does not change anymore and we may conclude that $f\left(x\right)$ is a polynomial of degree N. Otherwise, the degree of the Lagrange polynomial will continue to increase as $r-1$.

In summary, we have shown that the polynomial signature σ_p in (3) can be precisely reduced to the eigenvalue signature σ_λ in (1), provided that its length $r \geq N+1$ and that the set of different abscissa points $\left\{x_j\right\}_{1\leq j\leq r}$ is known. Since the abscissa points can be chosen freely, again we can construct infinitely many polynomial signatures σ_p, all equivalent, in the sense that they can reproduce the set of eigenvalues. The fact that the length r of the polynomial signature σ_p is tunable and not fixed at N as in the eigenvalue signature σ_λ will be exploited in the remainder.

2 Comparing graphs

The aim is to construct a method to compare two graphs $G_1\left(N_1, L_1\right)$ and $G_2\left(N_2, L_2\right)$ with possibly a different number of nodes, $N_1 \neq N_2$. In addition, we assume that the node labeling in G_1 is independent of that in G_2, otherwise the situation simplifies to computing the difference between two

matrices or two vectors with specified order in the row/columns or components. In that case, the lack of dimensions in G_2 due to a different $N_1 > N_2$, is mirrored by zero values in G_2 for those node labels of G_1 that are not contained in G_2. In other words, G_2 is augmented by an empty graph with $N_1 - N_2$ nodes so that both graphs possess N_1 nodes. This treatment is analogous to that in geometry and coordinate systems.

Since the eigenvalue signature σ_λ of a graph $G(N, L)$ has N vector components, the eigenvalue signature vectors $\sigma_\lambda(G_1)$ and $\sigma_\lambda(G_2)$ have possibly different dimensions. Linear algebra and geometry have established powerful concepts such as vector norms or projections to compare two vectors in a same dimensional space. Hence, in order to use these elegant vector comparison tools, the signature vectors of G_1 and G_2 should have the same length. The idea to compare two graphs $G(N_1, L_1)$ and $G(N_2, L_2)$ consists of constructing a polynomial signature σ_p with $r = \max(N_1, N_2)$ components. The discussion above has shown that the Lagrange polynomial constructed from the set $\{(x_j, c_{A_1}(x_j))\}_{1 \leq j \leq r}$ returns the exact characteristic polynomial corresponding to G_1 (and similarly $\{(y_j, c_{A_2}(y_j))\}_{1 \leq j \leq r}$ for G_2). The method is readily generalized to m graphs.

When comparing the signatures $\sigma_p(G_1)$ and $\sigma_p(G_2)$ of two graphs G_1 and G_2, respectively, we define a norm [1, Sec. 8.4] (because signatures are vectors) that measures their difference $\|\sigma_p(G_1) - \sigma_p(G_2)\|$. Apart from the uniqueness of a signature, another desirable property of a signature is that, if the graphs G_1 and G_2 do not differ "much", then $\|\sigma_p(G_1) - \sigma_p(G_2)\|$ should be "small" (and vice versa). In other words, when G_1 and G_2 are "close" in the graph space $\mathcal{G}$, a useful signature map should produce signatures $\sigma(G_1)$ and $\sigma(G_2)$ that are also "close" in the signature space $\mathcal{S}$.

We have shown in Section 1 that each unique signature σ_λ can be mapped into another unique signature $\sigma_{f(\lambda)}$, which implies that the norm can be made arbitrarily large. Hence, we propose to confine to signatures with unit norm, for example, by choosing the Euclidean norm $\sigma_p^T \sigma_p = \|\sigma_p\|_2^2 = 1$. Geometrically, this scaling means that each signature is a unit vector lying on the r-dimensional sphere and that the square of the norm of the difference

$$\begin{aligned} \|\sigma_p(G_1) - \sigma_p(G_2)\|_2^2 &= \left(\sigma_p^T(G_1) - \sigma_p^T(G_2)\right)\left(\sigma_p(G_1) - \sigma_p(G_2)\right) \\ &= 2\left(1 - \sigma_p^T(G_1)\,\sigma_p(G_2)\right) \leq 4 \end{aligned}$$

where the projection $\xi = \sigma_p^T(G_1)\,\sigma_p(G_2)$ of the two vectors equals the cosine of the angle between the vector $\sigma_p(G_1)$ and $\sigma_p(G_2)$. If $\xi = 1$, we may interpret the two graphs G_1 and G_2 as being the same, if $\xi = 0$, they can be regarded as orthogonal (or hardly possessing common features), while if $\xi = -1$, the two graphs are opposite.

2.1 Effect of the set of abscissa

It remains to investigate the effect of the set of abscissa $\{x_j\}_{1 \leq j \leq r}$ for G_1 and $\{y_j\}_{1 \leq j \leq r}$ for G_2 on the projection ξ. Invoking the definition (3) of a polynomial signature yields

$$\xi = \sigma_p^T(G_1)\,\sigma_p(G_2) = \frac{\sum_{j=1}^r c_{A_1}(x_j)\, c_{A_2}(y_j)}{\sqrt{\sum_{j=1}^r c_{A_1}^2(x_j) \sum_{j=1}^r c_{A_2}^2(y_j)}} \tag{4}$$

The denominator shows that we cannot choose all the zeros of either $c_{A_1}(x)$ or $c_{A_2}(x)$ as possible abscissa set. The sensitivity of the projection ξ to the sets of abscissa $\{x_j\}_{1 \leq j \leq r}$ and $\{y_j\}_{1 \leq j \leq r}$ is

represented by the vector Ξ with components $\frac{\partial \xi}{\partial x_l}$ and $\frac{\partial \xi}{\partial y_l}$ for $1 \leq l \leq r$. Ideally, we would like to have $\frac{\partial \xi}{\partial x_l} = \frac{\partial \xi}{\partial y_l} = 0$, implying that the choice of $\{x_j\}_{1 \leq j \leq r}$ and $\{y_j\}_{1 \leq j \leq r}$ does not influence ξ, or that we consider the maximum possible projection $\xi_{\max}$ as our distinguishing metric for the "similarity" between two graphs. Since $\frac{\partial \log \xi}{\partial x_l} = \frac{1}{\xi} \frac{\partial \xi}{\partial x_l}$, the solution of $\frac{\partial \log \xi}{\partial x_l} = 0$ is the same as that of $\frac{\partial \xi}{\partial x_l} = 0$ (and similarly for $\frac{\partial \xi}{\partial y_l}$) for any l and $\xi \neq 0$ where,

$$\begin{aligned}\frac{\partial \log \xi}{\partial x_l} &= \frac{\partial}{\partial x_l} \left\{ \log \sum_{j=1}^{r} c_{A_1}(x_j)\, c_{A_2}(y_j) - \frac{1}{2} \log \sum_{j=1}^{r} c_{A_1}^2(x_j) - \frac{1}{2} \log \sum_{j=1}^{r} c_{A_2}^2(y_j) \right\} \\ &= \frac{c'_{A_1}(x_l)\, c_{A_2}(y_l)}{\sum_{j=1}^{r} c_{A_1}(x_j)\, c_{A_2}(y_j)} - \frac{c_{A_1}(x_l)\, c'_{A_1}(x_l)}{\sum_{j=1}^{r} c_{A_1}^2(x_j)}\end{aligned}$$

Denoting the optimal solution of $\frac{\partial \log \xi}{\partial x_l} = 0$ by *, we obtain, for each $1 \leq l \leq r$, that either $c'_{A_1}(x_l^*) = 0$ and/or

$$\frac{c_{A_2}(y_l^*)}{\sum_{j=1}^{r} c_{A_1}\left(x_j^*\right) c_{A_2}\left(y_j^*\right)} = \frac{c_{A_1}(x_l^*)}{\sum_{j=1}^{r} c_{A_1}^2\left(x_j^*\right)} \tag{5}$$

Let us now assume that $r = N_1$. Since $c'_{A_1}(x)$ is a polynomial in x of degree $N_1 - 1$, at most $N_1 - 1$ *different* zeros of $c'_{A_1}(x)$, that interlace [1] with those of $c_{A_1}(x)$, can serve as optimal abscissa value and at least one x_k^* must be found by solving (5), which is a polynomial

$$c_{A_1}(x_k^*) = \frac{c_{A_2}(y_k^*) \sum_{j=1}^{r-1} c_{A_1}^2\left(x_j^*\right)}{\sum_{j=1}^{r-1} c_{A_1}\left(x_j^*\right) c_{A_2}\left(y_j^*\right)}$$

of degree r in x_k^*, given the right-hand side, thus all x_l^* for $1 \leq l < k \leq r$ and all $\left\{y_j^*\right\}_{1 \leq j \leq r}$. The optimal y_l^* values satisfy $c'_{A_2}(y_l^*) = 0$ and/or

$$\frac{c_{A_1}(x_l^*)}{\sum_{j=1}^{r} c_{A_1}\left(x_j^*\right) c_{A_2}\left(y_j^*\right)} = \frac{c_{A_2}(y_l^*)}{\sum_{j=1}^{r} c_{A_2}^2\left(y_j^*\right)} \tag{6}$$

Now, at most $N_2 - 1$ *different* zeros of $c_{A_2}(y_l)$ can be chosen as optimal abscissa and the $r - N_1 + 1$ other must be found from (6). After determining the zeros of $c'_{A_1}(x)$ and $c'_{A_2}(x)$, the remaining unknowns satisfy, for each l,

$$c_{A_2}(y_l^*) \sum_{j=1;j\neq l}^{r} c_{A_1}^2\left(x_j^*\right) - c_{A_1}(x_l^*) \sum_{j=1;j\neq l}^{r} c_{A_1}\left(x_j^*\right) c_{A_2}\left(y_j^*\right) = 0$$

and

$$c_{A_1}(x_l^*) \sum_{j=1;j\neq l}^{r} c_{A_2}^2\left(y_j^*\right) - c_{A_2}(y_l^*) \sum_{j=1;j\neq l}^{r} c_{A_1}\left(x_j^*\right) c_{A_2}\left(y_j^*\right) = 0$$

Elimination of $c_{A_2}(y_l^*)$ in the first and substitution in the second equation yields

$$c_{A_2}(y_l^*) = c_{A_1}(x_l^*) \frac{\sum_{j=1;j\neq l}^{r} c_{A_1}\left(x_j^*\right) c_{A_2}\left(y_j^*\right)}{\sum_{j=1;j\neq l}^{r} c_{A_1}^2\left(x_j^*\right)}$$

and

$$\sum_{j=1;j\neq l}^{r} c_{A_2}^2\left(y_j^*\right) \sum_{j=1;j\neq l}^{r} c_{A_1}^2\left(x_j^*\right) = \left(\sum_{j=1;j\neq l}^{r} c_{A_1}\left(x_j^*\right) c_{A_2}\left(y_j^*\right)\right)^2$$

By the Cauchy-Schwarz inequality [5], the last equation is only possible if $c_{A_1}\left(x_j^*\right) = c_{A_2}\left(y_j^*\right)$, for all $j \neq l$, leading to a maximal ξ. Thus, the remaining unknowns must satisfy $c_{A_2}\left(y_j^*\right) = c_{A_1}\left(x_j^*\right)$, after verification that all y_l^* are still different.

By enlarging the dimension N_2 of the eigenvalue signature $\sigma_\lambda(G_2)$ up to r, redundant information in the resulting polynomial signature $\sigma_p(G_2)$ is embedded, which by the "optimal" choice $c_{A_2}\left(y_j^*\right) = c_{A_1}\left(x_j^*\right)$ is somehow annihilated: the added $r - N_2$ dimensions do not lead to a differentiation in the corresponding vector components of $\sigma_p(G_1)$ and $\sigma_p(G_2)$. The analysis demonstrates that the vectors $\sigma_\lambda(G_1)$ and $\sigma_\lambda(G_2)$ can be used as well to compute the projection, resulting in a highest possible value

$$\begin{aligned}\xi_\lambda &= \frac{\sum_{j=1}^{N_2} \lambda_j(G_1)\,\lambda_j(G_2) + \sum_{j=N_2+1}^{N_1} \lambda_j^2(G_1)}{\sqrt{\left(\sum_{j=1}^{N_2} \lambda_j^2(G_2) + \sum_{j=N_2+1}^{N_1} \lambda_j^2(G_1)\right) \sum_{j=1}^{N_1} \lambda_j^2(G_1)}} \\ &= \frac{\sum_{j=1}^{N_2} \lambda_j(G_1)\,\lambda_j(G_2) + \sum_{j=N_2+1}^{N_1} \lambda_j^2(G_1)}{\sqrt{\left(2L_2 + \sum_{j=N_2+1}^{N_1} \lambda_j^2(G_1)\right) 2L_1}}\end{aligned}$$

because $\sum_{j=1}^{N} \lambda_j^2 = 2L$ (see [1, p. 30]). However, we can add $r - N_2$ components to a vector in many ways and each way may influence $\sum_{j=1}^{N_2} \lambda_j(G_1)\,\lambda_j(G_2)$. Thus, there appears to be too much arbitrariness which undermines the meaning and use of the projection ξ. Each vector component specifies one dimension in the r-dimensional space and, clearly, just as any coordinate system, the vectors $\sigma_p(G_1)$ and $\sigma_p(G_2)$ must have a consistent ordering in their components.

2.2 Revisiting the effect of abscissa

So far, we have introduced a polynomial signature σ_p for the eigenvector signature σ_λ because the vector length r of σ_p can be chosen arbitrarily long without loosing information (provided $r \geq N+1$). The graph comparison problem was formulated based on well-established concepts of linear algebra and geometry leading to the projection $\xi = \sigma_p^T(G_1)\,\sigma_p(G_2)$ of two vectors in the r-dimensional space. While the ordering of the vector components for the signature was not important, a consistent ordering of the coordinates between the two vectors $\sigma_p(G_1)$ and $\sigma_p(G_2)$ influences the scalar product ξ.

Within the proposed framework, a natural way to define an ordering in the vector components is by confining to one set of abscissa $\{x_j\}_{1\leq j\leq r}$ for both G_1 and G_2. Hence, the projection ξ in (4) reduces to

$$\xi = \sigma_p^T(G_1)\,\sigma_p(G_2) = \frac{\sum_{j=1}^{r} c_{A_1}(x_j)\,c_{A_2}(x_j)}{\sqrt{\sum_{j=1}^{r} c_{A_1}^2(x_j) \sum_{j=1}^{r} c_{A_2}^2(x_j)}}$$

and the equations for the optimal abscissa set $\left\{x_j^*\right\}_{1\leq j\leq r}$, satisfying $\frac{\partial \log \xi}{\partial x_l} = 0$ for $1 \leq l \leq r$, are

$$\frac{c'_{A_1}\left(x_l^*\right) c_{A_2}\left(x_l^*\right) + c'_{A_2}\left(x_l^*\right) c_{A_1}\left(x_l^*\right)}{\sum_{j=1}^{r} c_{A_1}\left(x_j^*\right) c_{A_2}\left(x_j^*\right)} - \frac{c_{A_1}\left(x_l^*\right) c'_{A_1}\left(x_l^*\right)}{\sum_{j=1}^{r} c_{A_1}^2\left(x_j^*\right)} - \frac{c_{A_2}\left(x_l^*\right) c'_{A_2}\left(x_l^*\right)}{\sum_{j=1}^{r} c_{A_2}^2\left(x_j^*\right)} = 0$$

from which the optimal projection follows, for any l, as

$$\xi^* = \frac{\left\{c'_{A_1}\left(x_l^*\right)c_{A_2}\left(x_l^*\right) + c'_{A_2}\left(x_l^*\right)c_{A_1}\left(x_l^*\right)\right\}\sqrt{\sum_{j=1}^r c_{A_1}^2\left(x_j^*\right)\sum_{j=1}^r c_{A_2}^2\left(x_j^*\right)}}{c_{A_1}\left(x_l^*\right)c'_{A_1}\left(x_l^*\right)\sum_{j=1}^r c_{A_2}^2\left(x_j^*\right) + c_{A_2}\left(x_l^*\right)c'_{A_2}\left(x_l^*\right)\sum_{j=1}^r c_{A_1}^2\left(x_j^*\right)}$$

The l equations form a set of l polynomial equations,

$$\begin{aligned}
0 = &\left\{c'_{A_1}\left(x_l^*\right)c_{A_2}\left(x_l^*\right) + c'_{A_2}\left(x_l^*\right)c_{A_1}\left(x_l^*\right)\right\}\sum_{j=1}^r c_{A_1}^2\left(x_j^*\right)\sum_{j=1}^r c_{A_2}^2\left(x_j^*\right)\\
&- c_{A_1}\left(x_l^*\right)c'_{A_1}\left(x_l^*\right)\sum_{j=1}^r c_{A_1}\left(x_j^*\right)c_{A_2}\left(x_j^*\right)\sum_{j=1}^r c_{A_2}^2\left(x_j^*\right)\\
&- c_{A_2}\left(x_l^*\right)c'_{A_2}\left(x_l^*\right)\sum_{j=1}^r c_{A_1}\left(x_j^*\right)c_{A_2}\left(x_j^*\right)\sum_{j=1}^r c_{A_1}^2\left(x_j^*\right)
\end{aligned}$$

Rewritten explicitly for $y = x_l^*$ after some tedious manipulations,

$$\begin{aligned}
0 = &\, c'_{A_1}(y)\left(c_{A_2}(y)\sum_{j=1;j\neq l}^r c_{A_1}^2\left(x_j^*\right) - c_{A_1}(y)\sum_{j=1;j\neq l}^r c_{A_1}\left(x_j^*\right)c_{A_2}\left(x_j^*\right)\right)\left(c_{A_2}^2(y) + \sum_{j=1;j\neq l}^r c_{A_2}^2\left(x_j^*\right)\right)\\
&+ c'_{A_2}(y)\left(c_{A_1}(y)\sum_{j=1;j\neq l}^r c_{A_2}^2\left(x_j^*\right) - c_{A_2}(y)\sum_{j=1;j\neq l}^r c_{A_1}\left(x_j^*\right)c_{A_2}\left(x_j^*\right)\right)\left(c_{A_1}^2(y) + \sum_{j=1;j\neq l}^r c_{A_1}^2\left(x_j^*\right)\right)
\end{aligned} \tag{7}$$

illustrates that y is a *real* zero of a polynomial of degree $4r-1$ in y (since all summations are independent of $y = x_l^*$). Each l of the r equations is similar, except that the summations are different per equation, which constitute the coupling between the r polynomial equations. The common zeros of both $c'_{A_1}(y)$ and $c'_{A_2}(y)$ satisfy the above polynomial equation. In general, solving the polynomial set of equations is involved. Moreover, it is difficult to demonstrate that there is an optimal solution $\{x_l^*\}_{1\leq l\leq r}$ for each pair $(c_{A_1}(x), c_{A_2}(x))$ of characteristic polynomials. However, when a solution $\{x_l^*\}_{1\leq l\leq r}$ is found, we may consider the resulting "best" projection ξ^* as a representative measure to compare the two graphs by one number.

In summary, observe again that we find a solution, leading to $\xi = 1$ and $c_{A_1}\left(x_l^*\right) = c_{A_2}\left(x_l^*\right)$ for all l, when we require that the brackets with differences in (7) are both set to zero. In that case, all x_l^* are the zeros of the difference polynomial $c_{A_1}(x) - c_{A_2}(x)$ that has degree r (unless $N_1 = N_2$), provided those zeros are distinct and real. The analysis here and the earlier conclusions indicate that this solution is not a desired one. On the other hand, the fact that we can find an optimal solution $\{x_l^*\}_{1\leq l\leq r}$ for each pair of graphs (G_1, G_2) with $\xi = 1$ (or at least very close to $\xi \to 1$) questions the possibility to find a meaningful number for graph comparison.

References

[1] P. Van Mieghem. *Graph Spectra for Complex Networks.* Cambridge University Press, Cambridge, U.K., 2011.

[2] W. Winterbach, P. Van Mieghem, M. Reinders, H. Wang, and D. de Ridder. Local topological signatures for network-based prediction of biological function. *Eighth IAPR International Conference on Pattern Recognition in Bioinformatics, Nice, France*, 17-20 June 2013.

[3] C. D. Godsil and B. D. McKay. Constructing cospectral graphs. *Aequationes Mathematicae*, 25:257–268, 1982.

[4] E. R. van Dam and W. H. Haemers. Which graphs are determined by their spectrum? *Linear Algebra and its Applications*, 373:241–272, 2003.

[5] P. Van Mieghem. *Performance Analysis of Communications Networks and Systems*. Cambridge University Press, Cambridge, U.K., 2006.

[6] E. C. Titchmarsh. *The Theory of Functions*. Oxford University Press, Amen House, London, 1964.

A Computation of $c_A(x)$ in (2) and entire functions

The computation of the polynomial signature σ_p may pose numerical problems[3]. When N is large, the product in (2) may not converge resulting in large overflows or underflows. Even some simple scaling is inadequate such as

$$c_A^*(x) = \frac{c_A(x)}{\prod_{j=1}^{N} \lambda_j} = \prod_{j=1}^{N} \left(1 - \frac{x}{\lambda_j}\right)$$

where[4] $\prod_{j=1}^{N} \lambda_j = \det A = c_A(0)$.

For large N, the product (2) does not necessary converge. This convergence problem is rather fundamental and related to the factorization of entire or integral functions. While a polynomial like $c_A(x)$ can be factored into the simple factors $(x - \lambda_i)$, less simple factors need to be considered in the factorization of entire functions.

This classical problem in analysis has been solved elegantly, mainly by Weierstrass and Hadamard, whose theory has given birth to the product form for the entire functions [6]. Entire functions $f(z)$ are analytic complex functions without singularities in the (finite) complex plane ($z \in \mathbb{C}$), but with an essential singularity at $z \to \infty$. Liouville has shown that an entire function without essential singularity at infinity is a constant. Entire functions can possess zeros in the finite plane. Examples of entire functions are $e^z, \sin z, \cos z$ and $\frac{1}{\Gamma(z)}$. Weierstrass has made a number of intriguing observations. First, the exponential function does not have zeros nor poles and consequently, when multiplying a function by an exponential function, the zeros are not altered. This means that the entire function

$$f(z) = \prod_{k=1}^{N} \left(1 - \frac{z}{\lambda_k}\right) e^{\sum_{j=1}^{m} \frac{z^j}{j\lambda_k^j}}$$

possesses precisely the same zeros as $c_A(z)$ for a finite integer m. The expressions

$$E(u;0) = 1 - u$$
$$E(u;m) = (1-u)\, e^{\sum_{j=1}^{m} \frac{u^j}{j}}$$

are called primary factors. Each primary factor vanishes when $u = 1$, but the behavior of $E(u;m)$ as $u \to 0$ depends on m. Indeed, for $|u| < 1$,

$$\log E(u;m) = \log(1-u) + \sum_{j=1}^{m} \frac{u^j}{j}$$

Using the Taylor series $\log(1-u) = -\sum_{j=1}^{\infty} \frac{u^j}{j}$, we obtain

$$\log E(u;m) = -\sum_{j=1+m}^{\infty} \frac{u^j}{j}$$

[3] Wynand Winterbach has conveyed this problem to me.

[4] We assume that the adjacency eigenvalues are different from zero, because we can always write $c_A(x) = \frac{x^m}{\prod_{j=1}^{N-m} \lambda_j} \prod_{j=1}^{N-m} \left(1 - \frac{x}{\lambda_j}\right)$, where the eigenvalue at zero has multiplicity m.

and

$$|\log E(u;m)| \leq \sum_{j=0}^{\infty} \frac{|u|^{j+m+1}}{j+m+1} \leq |u|^{m+1} \sum_{j=0}^{\infty} |u|^{j} = \frac{|u|^{m+1}}{1-|u|}$$

This inequality determines the convergence of a product of primary factors (see [6, p. 246-248]). For entire functions of finite order, we can determine the value of m precisely to assure convergence. An entire function $f(z)$ is of order ρ if $f(z) = O\left(e^{r^{\rho+\varepsilon}}\right)$ for everywhere $\varepsilon > 0$ when $|z| = r \to \infty$. For example, e^{z^k} is of order k, while e^{e^z} is of infinite order. In what follows, we shall suppose that $f(0)$ is not zero to simplify the analysis somewhat and a division of $f(z)$ by z^k does not affect the order. Titchmarsh [6, p. 249-250] proves that, if r_j is the modulus of the j-th zero of $f(z)$, then $\sum_{j=1}^{\infty} \frac{1}{r_j^{\alpha}}$ is convergent if $\alpha > \rho$. The lower bound $\alpha_{\min}$ of positive numbers α for which $\sum_{j=1}^{\infty} \frac{1}{r_j^{\alpha}}$ convergence is called the *exponent of convergence* of the zeros. Moreover, if $f(z)$ is of order ρ, then

$$f(z) = \prod_{j=1}^{\infty} E\left(\frac{z}{\lambda_j}, m\right)$$

and $\sum_{j=1}^{\infty} \frac{1}{r_j^{m+1}}$ is convergent, where the integer m, called the *genus* of the above canonical product, obeys $\rho - 1 \leq m \leq \alpha_{\min} \leq \rho$. A necessary and sufficient condition that $f(z) = \sum_{k=0}^{\infty} f_k z^k$ should be an entire function of finite order ρ is that

$$\lim_{k\to\infty} \frac{\log\left(1/|f_k|\right)}{k \log k} = \frac{1}{\rho} \tag{8}$$

The application of the theory of entire functions to the characteristic function $f(z) = \det(A - zI)$ of a graph when $N \to \infty$ still needs to be developed in graph theory. The lack of convergence of the product in (2) may limit the size N of the graphs. Because the entire functions are very likely not polynomials anymore, the "reconstructability" property of the Lagrange interpolation of polynomials disappears.