The number of walks and degree powers in directed graphs

Hanjo Täubig

Abstract

Fiol and Garriga proved that in undirected graphs the number \( w_k \) of walks of length \( k \) does not exceed the sum of the \( k \)-th powers of the vertex degrees, i.e., \( w_k \leq \sum_{x \in V} d(x)^k \). Here, we propose a generalization of this inequality for directed graphs using the geometric mean of the sums of the \( k \)-th powers of in- and out-degrees, namely, \( w_k^2 \leq (\sum_{x \in V} d_{in}(x))^k (\sum_{y \in V} d_{out}(y))^k) \). Further, we show that this inequality can be generalized for the case of nonnegative matrices, i.e., the sum of entries of the \( k \)-th matrix power is bounded from above by the geometric mean of the sums of the \( k \)-th powers of the row sums and column sums.

1 Introduction

Throughout the paper we assume that \( \mathbb{N} \) denotes the set of nonnegative integers. Let \( G = (V, E) \) be a directed graph having \( n \) vertices, \( m \) edges and adjacency matrix \( A \). We investigate (the number of) walks, i.e., sequences of vertices, where each pair of consecutive vertices \( v_i \) and \( v_{i+1} \) is connected by a directed edge \((v_i, v_{i+1}) \in E\). Nodes and edges can be used repeatedly in the same walk. The length \( k \) of a walk is counted in terms of edges.

For \( k \in \mathbb{N} \) and \( x, y \in V \), we denote by \( w_k(x, y) \) the number of walks of length \( k \) that start at vertex \( x \) and end at vertex \( y \). Since the graph is directed this number can be different from the number of walks of length \( k \) that start at vertex \( y \) and end at vertex \( x \). By \( s_k(x) = \sum_{y \in V} w_k(x, y) \) and \( e_k(x) = \sum_{y \in V} w_k(y, x) \) we denote the number of all walks of length \( k \) that start or end at node \( x \), resp. Consequently, \( w_k = \sum_{x \in V} s_k(x) = \sum_{x \in V} e_k(x) \) denotes the total number of walks of length \( k \). The set of all walks of length \( k \) is denoted by \( W_k \), i.e., \( w_k = |W_k| \). \( d_{in}(x) \) and \( d_{out}(x) \) denote the in-degree and the out-degree of vertex \( x \).

It is a well known fact that the \((i, j)\)-entry of \( A^k \) is the number of walks of length \( k \) that start at vertex \( i \) and end at vertex \( j \) (for all \( k \geq 0 \)). Fundamental observations about the number of walks are due to their decomposition into two or more segments:

**Observation 1.** For arbitrary graphs \( G = (V, E) \) and all vertices \( x, z \in V \) holds

\[
w_{k+\ell}(x, z) = \sum_{y \in V} w_k(x, y) \cdot w_{\ell}(y, z)
\]

and

\[
w_{k+p+\ell} = \sum_{(x^{\rightarrow} y) \in W_p} w_k(x) \cdot w_{\ell}(y)
\]

In particular, this implies:

\[
w_{k+1} = \sum_{x \in V} d_{in}(x) \cdot s_k(x) = \sum_{x \in V} d_{out}(x) \cdot e_k(x)
\]

\[
w_{k+\ell} = \sum_{x \in V} e_k(x) \cdot s_{\ell}(x) = \sum_{x \in V} e_{\ell}(x) \cdot s_k(x)
\]

*Institut für Informatik, Technische Universität München, D-85748 Garching, Germany, taeubig@in.tum.de*
2 Walks and degree powers

The following inequality for undirected graphs was conjectured by Marc Noy and proven by Fiol and Garriga [FG09]:

**Theorem 2.** In any undirected graph, the number \( w_k \) of walks of length \( k \) does not exceed the sum of the \( k \)-th powers of the vertex degrees, i.e.,

\[
    w_k \leq \sum_{x \in V} d_x^k.
\]

In the following, we discuss possible generalizations of this theorem to directed graphs. The conceivable inequality \( w_k \leq \sum_{x \in V} d_{in}(x)^k \) is invalid. For instance, it is violated by the graph shown in Figure 1. Because of the reversely directed counterpart of this graph, the same applies to the inequality \( w_k \leq \sum_{x \in V} d_{out}(x)^k \). Also, trying to generalize the inequality by using direct products of \( d_{in}(x) \) and \( d_{out}(x) \) is not successful, since, e.g., \( w_k \leq \sum_{x \in V} \sqrt{d_{in}(x) \cdot d_{out}(x)}^k \) is violated for \( k = 1 \) by the graph consisting of only one directed edge.

**Observation 3.** The following inequalities are invalid generalizations of Theorem 2:

\[
    w_k \not\leq \sum_{x \in V} d_{in}(x)^k \\
    w_k \not\leq \sum_{x \in V} d_{out}(x)^k \\
    w_k \not\leq \sum_{x \in V} \sqrt{d_{in}(x) \cdot d_{out}(x)}^k
\]

While the power sum for \( d_{in}(x) \) or \( d_{out}(x) \) alone is not suitable for bounding \( w_k \), we will show that a combination (namely, the geometric mean) of both sums is sufficient. To this end, we first show that for the consideration of power sums with exponent \( q \) over the set of walks of length \( p \) the total cannot decrease if we shorten the walk length while at the same time the exponent is increased by the same difference.

**Lemma 4.** For every directed graph \( G = (V, E) \) and for all nonnegative integers \( p, q \in \mathbb{N} \) holds

\[
    \left( \sum_{(x \rightarrow y) \in W_p} d_{in}(x)^q \right) \left( \sum_{(x \rightarrow y) \in W_p} d_{out}(y)^q \right) \leq \left( \sum_{(x \rightarrow w \rightarrow y) \in W_{p-1}} d_{in}(x)^{q+1} \right) \left( \sum_{(x \rightarrow w \rightarrow y) \in W_{p-1}} d_{out}(y)^{q+1} \right)
\]

**Proof.** The proof starts with decomposing and counting walks of length \( p \) from \( x \) to \( y \), denoted by \( (x \rightarrow y) \), into walks of length \( p - 1 \) which is prepended or followed by a single edge, i.e., \( (x \rightarrow w \rightarrow y) \) and \( (x \rightarrow z \rightarrow y) \), resp.
Theorem 5. For every directed graph $G = (V, E)$ and for all nonnegative integers $p \in \mathbb{N}$ holds
\[
w_p^2 \leq \left( \sum_{x \in V} d_{in}(x)^p \right) \left( \sum_{y \in V} d_{out}(y)^p \right)
\]
Lemma 7. For every nonnegative $w \in \mathbb{R}$, therefore, at least one of the two power sums must be greater than or equal to $w$.

Proof. The proof works by repeatedly applying Lemma 4 to $w_p^2$:

$$w_p^2 = \left( \sum_{x \in V} \sum_{y \in V} w_p(x, y) d_{in}(x)^{2p} \right) \left( \sum_{x \in V} \sum_{y \in V} w_p(x, y) d_{out}(y)^{2p} \right)$$

$$\leq \left( \sum_{x \in V} \sum_{w \in V} w_{p-1}(x, w) d_{in}(x)^{p} \right) \left( \sum_{z \in V} \sum_{y \in V} w_{p-1}(z, y) d_{out}(y)^{p} \right)$$

$$\vdots$$

$$\leq \left( \sum_{x \in V} \sum_{w \in V} w_0(x, w) d_{in}(x)^p \right) \left( \sum_{z \in V} \sum_{y \in V} w_0(z, y) d_{out}(y)^p \right) = \left( \sum_{x \in V} d_{in}(x)^p \right) \left( \sum_{y \in V} d_{out}(y)^p \right)$$

The last equality follows from the fact that $w_0(x, y)$ is 1 for $x = y$ and 0 otherwise.

This means, although $w_k \not\leq \sum_{x \in V} d_{in}(x)^k$ and $w_k \not\leq \sum_{x \in V} d_{out}(x)^k$, we know for the geometric mean of the two power sums that

$$w_k \leq \sqrt{\left( \sum_{x \in V} d_{in}(x)^k \right) \left( \sum_{x \in V} d_{out}(x)^k \right)}.$$ 

Therefore, at least one of the two power sums must be greater than or equal to $w_k$:

$$w_k \leq \max \left\{ \sum_{x \in V} d_{in}(x)^k, \sum_{x \in V} d_{out}(x)^k \right\}$$

and of course, the inequality of arithmetic and geometric means implies

$$w_k \leq \frac{1}{2} \left( \sum_{x \in V} d_{in}(x)^k + d_{out}(x)^k \right).$$

Note that Theorem 5 contains Theorem 2 by Fiol and Garriga as a special case ($d_{in}(x) = d_{out}(x)$).

### 3 Nonnegative Matrices

For an arbitrary $n \times n$-matrix $A$, let $\text{sum}(A)$ denote the sum of the entries of $A$. The set of matrix indices $\{1, \ldots, n\}$ is denoted by $[n]$. Further, we define $a_{ij}^{[p]}$ to be the $(i, j)$-entry of $A^p$. The row and column sums shall be denoted by $r_i$ and $c_j$ ($i, j \in [n]$).

Actually, Theorem 2 is only the special case for adjacency matrices of the following theorem (see Corollary (3.24) in the book by Berman and Plemmons [BP94]) that holds for powers of symmetric matrices and their row or column sums:

**Theorem 6.** For every symmetric matrix with row sums $r_i$ ($i \in [n]$) holds:

$$\text{sum} (A^k) \leq \sum_{i=1}^n r_i^k$$

Now, we will generalize this theorem to the case of arbitrary nonnegative matrices.

**Lemma 7.** For every nonnegative $n \times n$-matrix $A = (a_{ij})$ with row sums $r_i$ and column sums $c_i$ ($i \in [n]$) holds:

$$\left( \sum_{x \in [n]} \sum_{y \in [n]} a_{xy}^{[p]} c_x^q \right) \left( \sum_{x \in [n]} \sum_{y \in [n]} a_{xy}^{[p]} c_y^q \right) \leq \left( \sum_{x \in [n]} \sum_{y \in [n]} a_{xy}^{[p-1]} c_x^{q+1} \right) \left( \sum_{x \in [n]} \sum_{y \in [n]} a_{xy}^{[p-1]} c_y^{q+1} \right)$$

4
Proof.

\[
\left( \sum_{x \in [n]} \sum_{y \in [n]} a_{x,y}^p \right) \left( \sum_{x \in [n]} \sum_{y \in [n]} a_{x,y}^p \right) = \left( \sum_{x,y \in [n]} a_{x,y}^{p-1} \right) \left( \sum_{x,y \in [n]} a_{x,y}^{p-1} \right) = \left( \sum_{x \in [n]} \sum_{y \in [n]} a_{x,y}^{p-1} c_2^x \right) \left( \sum_{y \in [n]} \sum_{x \in [n]} a_{x,y}^{p-1} c_2^y \right) = \left( \sum_{x \in [n]} \sum_{y \in [n]} a_{x,y}^{p-1} c_x \right) \left( \sum_{y \in [n]} \sum_{x \in [n]} a_{x,y}^{p-1} c_y \right) \]

Theorem 8. For every nonnegative \( n \times n \)-matrix \( A = (a_{ij}) \) and \( p \in \mathbb{N} \) holds

\[
\text{(sum}(A^p)\text{)}^2 \leq \left( \sum_{x \in [n]} c_x^p \right) \left( \sum_{y \in [n]} r_y^p \right)
\]

Proof. The proof works by repeatedly applying Lemma 7 to the squared entry sum of matrix \( A^p \):

\[
\text{(sum}(A^p)\text{)}^2 = \left( \sum_{x \in [n]} \sum_{y \in [n]} a_{x,y}^{p-1} c_x^1 \right) \left( \sum_{y \in [n]} \sum_{x \in [n]} a_{x,y}^{p-1} r_y^1 \right) \leq \left( \sum_{x \in [n]} \sum_{y \in [n]} a_{x,y}^{p-1} c_x^1 \right) \left( \sum_{y \in [n]} \sum_{x \in [n]} a_{x,y}^{p-1} r_y^1 \right) \leq \left( \sum_{x \in [n]} \sum_{y \in [n]} a_{x,y}^{p-1} c_x^1 \right) \left( \sum_{y \in [n]} \sum_{x \in [n]} a_{x,y}^{p-1} r_y^1 \right) = \left( \sum_{x \in [n]} c_x^p \right) \left( \sum_{y \in [n]} r_y^p \right)
\]

The last equality follows from the fact that \( a_{ij}^{(0)} \) is 1 for \( i = j \) and 0 otherwise, since \( A^0 \) is the identity matrix. \( \square \)
The last theorem implies an even more general form of Theorem 5 for walks:

**Corollary 9.** For every directed graph $G = (V, E)$ and for all nonnegative integers $p \in \mathbb{N}$ holds:

$$w_{pk}^2 \leq \left( \sum_{x \in V} e_k(x)^p \right) \left( \sum_{y \in V} s_k(y)^p \right)$$

**References**
