PageRank: Matrix Formulation

- **Stochastic adjacency matrix $\boldsymbol{M}$**
  - Let page $\boldsymbol{i}$ has $\boldsymbol{d_i}$ out-links
  - If $\boldsymbol{i} \rightarrow \boldsymbol{j}$, then $\boldsymbol{M_{ji}} = \frac{\boldsymbol{1}}{\boldsymbol{d_i}}$ else $\boldsymbol{M_{ji}} = \boldsymbol{0}$
    - $\boldsymbol{M}$ is a **column stochastic matrix**
      - Columns sum to **1**
- **Rank vector $\boldsymbol{r}$:** vector with an entry per page
  - $\boldsymbol{r_i}$ is the importance score of page $\boldsymbol{i}$
  - $\sum_i \boldsymbol{r_i} = \boldsymbol{1}$
- **The flow equations can be written**

$$r_j = \sum_{i \rightarrow j} \frac{r_i}{\mathrm{d_i}}$$

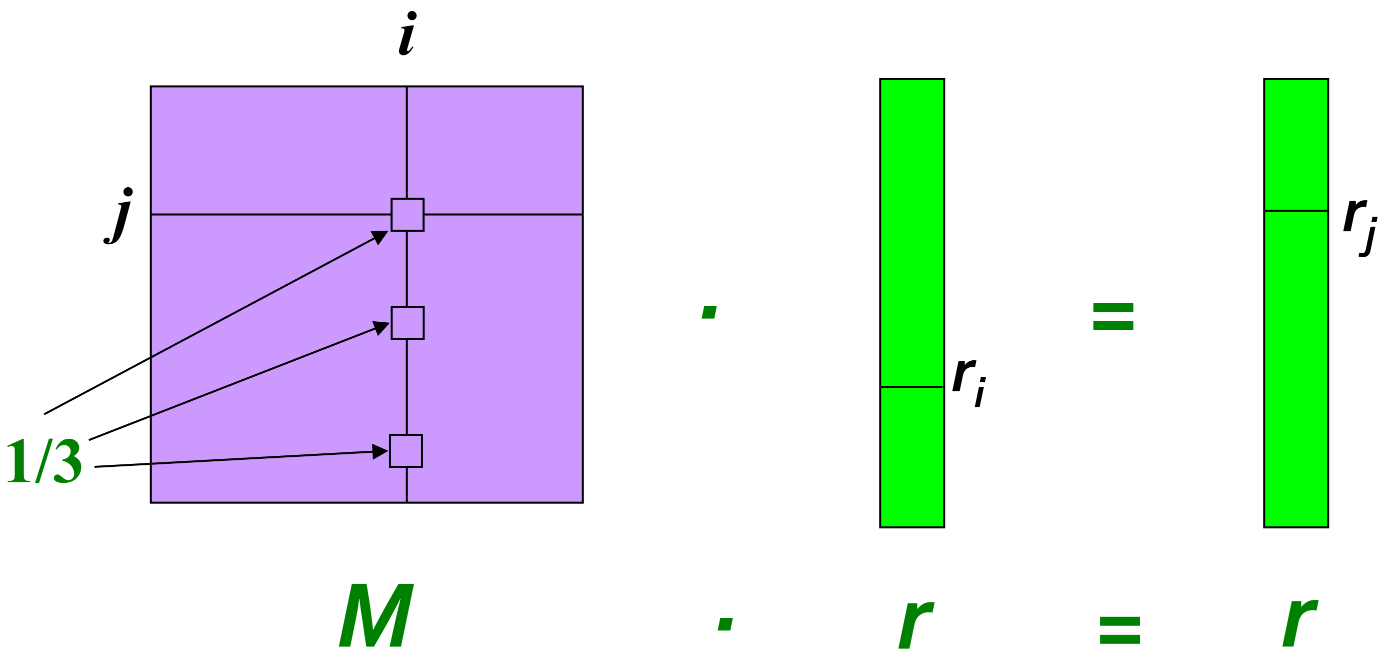
$$\boldsymbol{r} = \boldsymbol{M} \cdot \boldsymbol{r}$$

Example

- Remember the flow equation: $r_j = \sum_{i \to j} \frac{r_i}{\mathrm{d_i}}$
- Flow equation in the matrix form

$$\boldsymbol{M} \cdot \boldsymbol{r} = \boldsymbol{r}$$

  - Suppose page *i* links to 3 pages, including *j*



Eigenvector Formulation

NOTE: $\boldsymbol{x}$ is an eigenvector with the corresponding eigenvalue **λ** if:

$$Ax = \lambda x$$

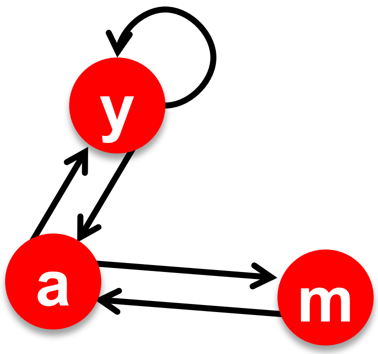
- The flow equations can be written

$$r = M \cdot r$$

- So the **rank vector** $\boldsymbol{r}$ is an **eigenvector** of the stochastic web matrix $\boldsymbol{M}$
  - In fact, its first or principal eigenvector, with corresponding eigenvalue ***1***
    - Largest eigenvalue of $\boldsymbol{M}$ is **1** since $\boldsymbol{M}$ is column stochastic
      - ***Why?*** *We know* **r** *is unit length and each column of* **M** *sums to one, so* $\boldsymbol{Mr} \leq \mathbf{1}$

- **We can now efficiently solve for *r*!**
  The method is called Power iteration

Example: Flow Equations & M



$$
\begin{array}{c|ccc}
 & y & a & m \\
\hline
y & \tfrac{1}{2} & \tfrac{1}{2} & 0 \\
a & \tfrac{1}{2} & 0 & 1 \\
m & 0 & \tfrac{1}{2} & 0
\end{array}
$$

$$r = M \cdot r$$

$$
\begin{aligned}
r_y &= r_y/2 + r_a/2 \\
r_a &= r_y/2 + r_m \\
r_m &= r_a/2
\end{aligned}
$$

$$
\begin{bmatrix} y \\ a \\ m \end{bmatrix} = \begin{bmatrix} \tfrac{1}{2} & \tfrac{1}{2} & 0 \\ \tfrac{1}{2} & 0 & 1 \\ 0 & \tfrac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} y \\ a \\ m \end{bmatrix}
$$