

数理科学科だより

ググりたくなる線形代数

今回は「ググる」という言葉の説明から始め、線形代数が身近に応用されている例についてご紹介したいと思います。

・「ググる」とは

「ググる」とは検索エンジン Google を使用して検索することです。また英語圏においても「google」は「Google で検索する」という意味の動詞として用いられています。このような言葉が誕生するに至る程、Google が検索エンジンの中で大きなシェアを占めることに成功できたのはなぜなのでしょう。その理由の一つは検索結果の正確さであり、実はその陰には線形代数で学ぶ固有方程式があるのです。

・ Google 検索エンジンを支えるページランクの考え方

私たちが検索エンジンに求めることは、ニーズにあった“重要なページ”を探し当てることかと思えます。そこで Google は“重要なページ”を計る指標としてページランク(名前の由来はウェブページのページと Google 創設者の一人 ラリー・ペイジの姓をかけたものです)という概念を導入しました。ウェブページ i のページランク r_i を

$$r_i = \sum_{j \in A_i} \frac{r_j}{|j|}$$

として定めます。ここで、 A_i はウェブページ i にリンクしているウェブページの集合、 $|j|$ はウェブページ j から出ていくリンクの数を表すとして、ページランクの値が大きい程、“重要なページ”とみなします。

さて、このページランクはどのようにして求めるのでしょうか。ページランク $r = (r_1, \dots, r_n)$ は行列を用いて $r = rA$ という形で表されます。従って、ページランクを計算するためには行列 A の固有値 1 に対応する固有ベクトルを求めれば良いことがわかります。因みに全ての成分が正の実数である正方行列には実数の固有値が存在し、固有値に対応する固有ベクトルの成分もまた実数となる(ペロン=フロベニウスの定理)ことが知られています。これを応用することで、各ウェブページにページランクを実数として(スカラー倍を除いて)唯一つ与えることができます。従って、ページランクを用いてウェブページの“重要さ”を比較することができます。このようにページランクの考え方は比較的シンプルですが、それが上手く“機能”する理由の一つが線形代数で学ぶ固有方程式に関する理論なのです。

今回はご紹介することができませんでしたが、他にもページランクについては興味深いお話が沢山あります。しかしながら、全てをここに記すには余白が狭すぎるようです。もし興味をお持ちの方が居られましたら、まずはググってみてください(文：塚本)。

編集：山口大学理学部数理科学科

連絡先：083-933-5210 (理学部学務係)

<http://www.sci.yamaguchi-u.ac.jp/ja/dep/math/info/news.html>