Adjacency Matrices

Chapter 2 - Project B

Varun Chitturi, Aditya Patel

1 Introduction

The purpose of this project is to show how powers of a matrix may be used to investigate graphs. smoke trees.

It is worthy of exploring it.

2 Concepts

The general form of Nesterov's accelerated gradient method to minimize the cost function $f(x)$ can be written in the following form:

$$x^{k+1} = x^k - \gamma \nabla f(x^k + \mu(x^k - x^{k-1})) + \mu(x^k - x^{k-1}).$$

The parameters γ and μ are difficult to tune when $f(x)$ is non-convex.

3 Anderson acceleration

We apply Anderson acceleration to the problem $x = x - \alpha \nabla f(x)$:

$$x^{k+1} = x^k - \alpha \nabla f(x^k) + (1-\alpha_1)\alpha[\nabla f(x^k) - \nabla f(x^{k-1})] + (1-\alpha_1)(x^{k-1} - x^k)$$

where $\alpha_1 = \arg\min_{\alpha_1} \|\alpha_1 \nabla f(x^k) + (1-\alpha_1)\nabla f(x^{k-1})\|_2^2 = -\frac{\|\nabla f(x^k) - \nabla f(x^{k-1})\|_2^2}{\langle \nabla f(x^k) - \nabla f(x^{k-1}), \nabla f(x^{k-1}) \rangle}$.

There is a difference of previous iteration in Anderson acceleration as well as Nesterov's accelerated gradient method. It seems that

$$\gamma \nabla f(x^k + \mu(x^k - x^{k-1})) \approx (1-\alpha_1)\alpha[\nabla f(x^k) - \nabla f(x^{k-1})].$$

And what is the connection of these two methods in mathematics?