# Nonlinear eigenvalue problems

Shota Uka

27. August 2025

CAMBRIDGE IMAGE ANALYSIS
DAMTP - UNIVERSITY OF CAMBRIDGE

# Contents

# 1. Introduction – nonlinear eigenvalue problems

We want to generalise the notion of an eigenvalue and eigenvector to the nonlinear case. Let us first consider the linear case.
Let $A \in \mathbb{R}^{n \times n}$. We remember that $u \in \mathbb{R}^n \setminus \{0\}$ is an eigenvector of $A$ iff $Au = \lambda u$ for $\lambda \in \mathbb{R}$.

# 1. Introduction – nonlinear eigenvalue problems

$$Au = \lambda u,\ \ u \neq 0, A \text{ symmetric}$$

$$\Updownarrow$$

$$u = \text{stationary point of } R(u) = \frac{\langle Au, u\rangle}{\langle u, u\rangle}$$

$$\max_{x\in\mathbb{R}^n\setminus\{0\}} \frac{\langle Ax, x\rangle}{||x||^2} = \lambda_{\max}$$

$$\min_{x\in\mathbb{R}^n\setminus\{0\}} \frac{\langle Ax, x\rangle}{||x||^2} = \lambda_{\min}.$$

# 1. Introduction – nonlinear eigenvalue problems

$Au = \lambda Bu, \; u \neq 0, A$ symmetric, $B$ symmetric, positive definite

$$\Updownarrow$$

$$u = \text{stationary point of } R(u) = \frac{\langle Au, u\rangle}{\langle Bu, u\rangle}$$

$$\max_{x\in\mathbb{R}^n\setminus\{0\}} \frac{\langle Ax, x\rangle}{\langle Bx, x\rangle} = \lambda_{\max}$$

$$\min_{x\in\mathbb{R}^n\setminus\{0\}} \frac{\langle Ax, x\rangle}{\langle Bx, x\rangle} = \lambda_{\min}$$

# 1. Introduction – nonlinear eigenvalue problems

Let $F, G : \mathbb{R}^n \to \mathbb{R}$ (such that $F, G$ satisfy suitable conditions). We define the eigenvectors of $(F, G)$ as the stationary points of the generalised Rayleigh quotient

$$\lambda(u) = \frac{F(u)}{G(u)}$$

where $G > 0$, i.e. $u$ satisfies the following generalised eigenvalue problem

$$\nabla F(u) = \lambda(u) \nabla G(u).$$

# 1. Introduction – the PM

## Definition (Power Method)

Given an initial vector $u_0$ with $\|u_0\| = 1$, the power method generates a sequence $(u_k)_{k\in\mathbb{N}}$ via the iteration:

$$
\begin{aligned}
v_{k+1} &= Au_k, \\
u_{k+1} &= \frac{v_{k+1}}{\|v_{k+1}\|}.
\end{aligned}
$$

Under suitable conditions on $A$ and $u_0$, the sequence $(u_k)_{k\in\mathbb{N}}$ converges to an eigenvector corresponding to $\lambda_{\max}$, and the sequence of Rayleigh quotients $(R(u_k))_{k\in\mathbb{N}}$ converges to $\lambda_{\max}$. The inverse power method computes the eigenvector corresponding to the smallest eigenvalue.

# 1. Introduction – the GRPM

**Definition (Generalised Regular Power Method)**

The sequence $(u_k, \lambda_k)_{k\in\mathbb{N}}$ is generated by the following scheme:

$$u_k = \arg\max_u \left\{ \langle \nabla F(u_{k-1}), u \rangle - \lambda_{k-1} G(u) - D(u, u_{k-1}) \right\}, \tag{1}$$

$$\lambda_k = \frac{F(u_k)}{G(u_k)}, \tag{2}$$

where $D(u, v) \geq 0$ is a distance term (e.g., the squared Euclidean distance) with $D(u, u) = 0$.

# 1. Introduction – the GIPM

**Definition (Generalised Inverse Power Method)**

The sequence $(u_k, \lambda_k)_{k\in\mathbb{N}}$ is generated by:

$$u_k = \arg\min_u \left\{ F(u) - \lambda_{k-1} \langle \nabla G(u_{k-1}), u \rangle + D(u, u_{k-1}) \right\}, \qquad (3)$$

$$\lambda_k = \frac{F(u_k)}{G(u_k)}. \qquad (4)$$

# 2. The KL property – definition

## Definition

Let $\eta > 0$. By $\mathcal{K}_n$ we denote a class of functions $\phi : [0, \eta) \to \mathbb{R}_+$ such that

1. **Continuity:** $\phi$ is continuous on $[0, \eta)$,
2. **Concavity:** $\phi$ is concave on $[0, \eta)$,
3. **Origin:** $\phi(0) = 0$,
4. **Strictly increasing:** $\phi \in C^1((0, \eta))$ with $\phi' > 0$ (i.e. $\phi$ is strictly increasing).

In the following we will always consider a function $f$ such that $f \in C^1(\mathbb{R}^n)$.

# 2. The KL property – definition

## Definition

1. A function $f : \mathbb{R}^n \to \mathbb{R}$ satisfies the Kurdyka-Lojasiewiez property (KL property) at $\bar{x}$ if there exists a neighbourhood $U$ of $\bar{x}$, $\eta > 0$ and $\phi : [0, \eta) \to \mathbb{R}_+$ such that $\phi \in \mathcal{K}_n$ and

$$\phi^{'}(f(x) - f(\bar{x}))||\nabla f(x)|| \geq 1$$

   for all $x \in U$ such that $f(\bar{x}) < f(x) < f(\bar{x}) + \eta$. If such as function $\phi$ exists, we call it the deregulariser of $f$ at $\bar{x}$.
2. If a function $f : \mathbb{R}^n \to \mathbb{R}$ satisfies the KL-property at every $x \in \mathbb{R}^n$, then we call $f$ a KL-function.

# 2. The KL property – definition

We want to analyse the KL-property now: The KL property is in some sense a measure of whether the function is amenable to sharpness around a point or not. A function $f : \mathbb{R}^n \to \mathbb{R}$ is sharp around $\bar{x}$ iff there is a neighbourbood $U$ of $\bar{x}$ such that

$$||\nabla f(x)|| \geq 1$$

for all $x \in U \setminus \{\bar{x}\}$. The KL property at $\bar{x}$ means therefore that $f$ can be sharpened around $\bar{x}$ by reparameterizing its values with a function $\phi \in \mathcal{K}_n$.

# 2. The KL property – example

The KL property is always satisfied at non-stationary points, but it can also be satisfied at stationary points, as the following example shows:

**Example**

Let $f(x) := x^2$. Then $f$ has the KL property at 0 with respect to $U = \mathbb{R}, \eta = \infty$ and $\phi(t) = \sqrt{t}$.

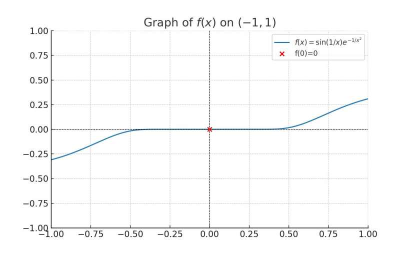
# 2. The KL property – example

If the function is too flat around a critical point, then it won't satisfy the KL property there:

## Example

Let

$$f(x) = \begin{cases} \sin(\frac{1}{x})e^{-\frac{1}{x^2}}, & x \neq 0, \\ 0, & x = 0. \end{cases}$$




Then $f$ fails to satisfy the KL property at 0.

# 2. The KL property – convexity

We want to explore the relation of the KL property to convexity. Convex functions are not necessarily KL. However uniform convexity suffices:

**Definition**

Let $f : \mathbb{R}^n \to \mathbb{R}$. $f$ is *uniformly convex* with modulus $\varphi$ if for every $x, y \in \mathbb{R}^n$ and every $\lambda \in (0, 1)$ we have

$$f\big(\lambda x + (1-\lambda)y\big) + \lambda(1-\lambda)\,\varphi\big(\|x-y\|\big) \leq \lambda f(x) + (1-\lambda)f(y), \quad (5)$$

where $\varphi : \mathbb{R}_+ \to [0, \infty]$ is strictly increasing and $\varphi(t) = 0$ iff $t = 0$.

# 2. The KL property – convexity

**Theorem**

*Let $f : \mathbb{R}^n \to \mathbb{R}$ be uniformly convex with modulus $\phi = c||\cdot||^r$ for $r \geq 1$. Then $f$ has the KL-property on $\mathbb{R}^n$ with respect to $U = \mathbb{R}^n, \eta = \infty$ and $\phi(t) = rc^{-\frac{1}{r}} t^{\frac{1}{r}}$.*

If $n = 1$, we have:

**Theorem**

*Every convex $C^1$ function $f : \mathbb{R} \to \mathbb{R}$ is KL.*

# 3. Convergence Analysis of the GIPM

**Theorem**

*Let $F, G : \mathbb{R}^n \to \mathbb{R}$ such that $F, G \in C^1(\mathbb{R}^n)$, $G(u) > 0$ for all $u \in \mathbb{R}^n$ and $\nabla G$ Lipschitz-continuous. Let the sequence $(u_k, \lambda_k)_{k\in\mathbb{N}}$ be generated by (3)-(4) and let $\lambda$ be bounded from below. Assume that $\lambda$ satisfies the KL-property at every critical point. Assume moreover that the distance function $D$ is given by*

$$D(u, v) = \frac{1}{2\gamma}||u - v||^2,$$

*where $\gamma > 0$ such that*

$$\frac{1}{2L_G \max\limits_{k\in\mathbb{N}} \lambda_k} > \gamma$$

*and that $(u_k)_{k\in\mathbb{N}}$ is bounded. Then the sequence $(u_k)_{k\in\mathbb{N}}$ converges to a critical point of $\lambda$.*

# Thank you!

I would like to thank my supervisors Carola-Bibiane Schönlieb, Martin Benning and Matthias Ehrhardt. In particular, I am deeply grateful to Carola-Bibiane Schönlieb for her invaluable advice and support not only throughout the project but also since we first met.
I also wish to thank Davide Murari for introducing me to PyTorch. Many thanks to my colleagues from the internship, who made this experience both inspiring and enjoyable, and also to Emma Beniston for being a wonderful mentor.
I would also like to thank the organisers of the programme, especially Sally, for always being so helpful and kind. Finally, I wish to thank all those whose generous donations made this experience possible.
This project was funded by the Philippa Fawcett Internship Programme.