Rational approximation, its role in different branches of mathematics and applications

Nadezda Sukhorukova

Swinburne University of Technology

nsukhorukova@swin.edu.au

November 3, 2021

Overview

1. Origins and motivation
   - Chebyshev approximation
   - Rational approximation
2. Approximation and Optimisation
   - Formulation and methods
   - Quasiconvexity
   - Methods
3. Applications
   - Computational Chemistry
   - Deep learning

Chebyshev approximation

In this talk, we are working with uniform (Chebyshev) approximation:

$$\min_{A} \max_{t \in Q} |f(t) - g(A, t)|,$$

where A are the decision variables.
The optimality conditions are based on maximal deviation points. In the case of univariate function approximation, the conditions are based on the notion of of alternating sequence.
In the case of univariate polynomial approximation:

$$\min_{A=(a_0,\ldots,a_n)} \max_{t \in [a,b]} |f(t) - (a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n)|.$$

Theorem

(Chebyshev) The solution is optimal if and only there are $n+2$ alternating points.

Multivariate Chebyshev approximation

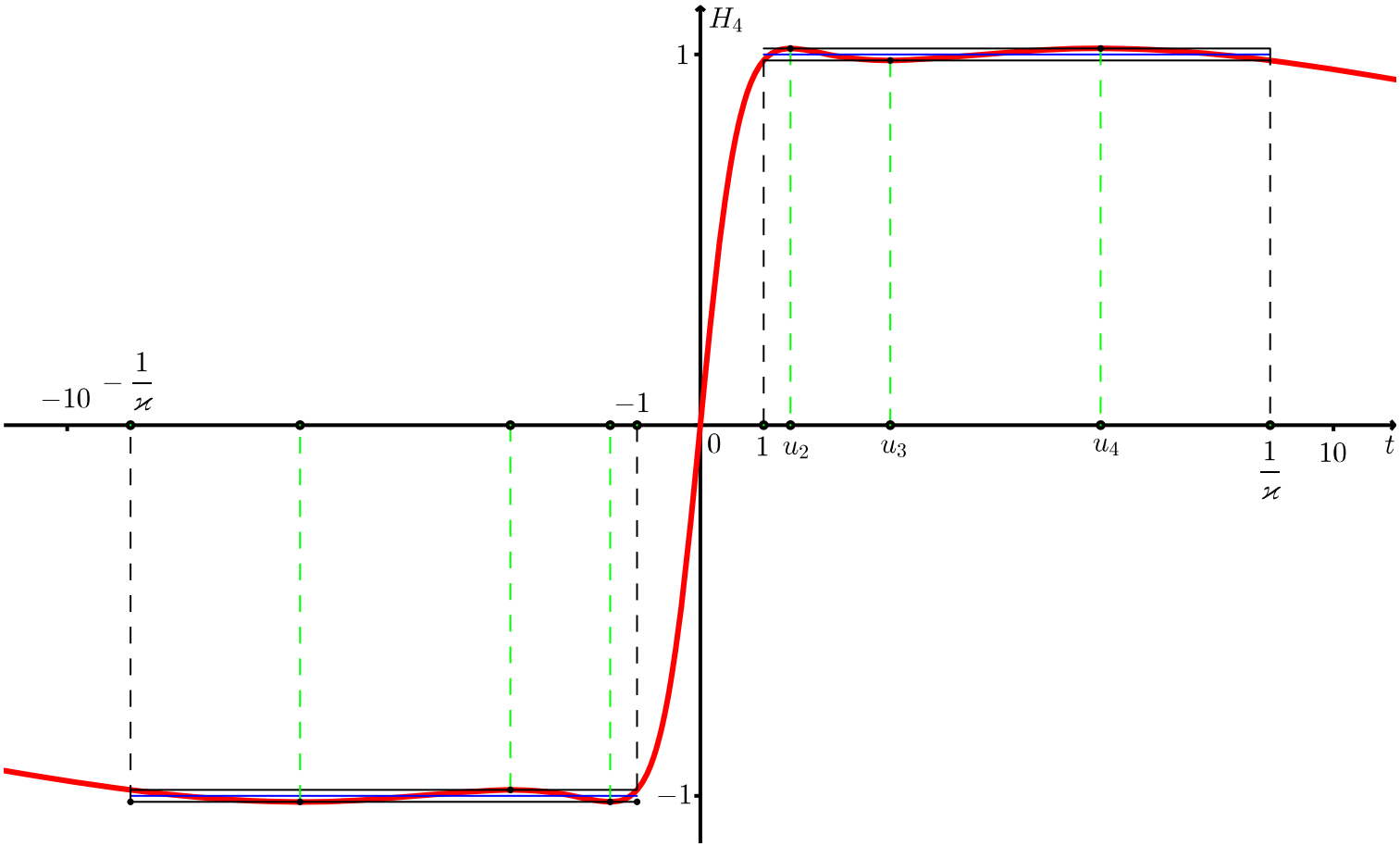
In the case of multivariate function approximation, the points are not totally ordered and therefore the extension is problematic, but there are some results: more complex geometrical constructions are involved.
In some papers the authors reduce multivariate alternating sequence (alternance) to the alternation of certain determinants.
The main computational problem here is that multivariate monomials do NOT form a Chebyshev system.

From polynomial to rational approximation.

Polynomial approximation (approximation by a linear combination of monomials). The coefficients are subject to optimisation. Most results can be extended to the case when the basis functions are not restricted to monomials, but form a Chebyshev system.

- Polynomial approximation (approximation by a linear combination of monomials). The coefficients are the decision variables.
- Polynomial splines or piecewise polynomials (fixed or free knots: points of switching from one polynomial to another).
- Rational approximation (approximation by a ratio of two polynomials).
- Generalised rational approximation (ratio of linear forms, the basis functions are not limited to monomials).

Approximation of $sign(x)$: the degree of the polynomials is 4. The picture was produced by Grigoriy Tamasyan.



Free knot spline approximation and rational approximation are related

A number of studies (Petrushev and Popov [12]) indicate that rational functions may be a suitable substitution for free knot polynomial spline approximation whose corresponding optimisation problems are very complex and there is no efficient computational tool for constructing the corresponding approximations (Nurnberger et. al. [4]). In particular, this problem was listed as one of the most important open problem in approximation [4].

Approximation

Consider uniform approximation by a rational function $R_{nk}(t)$, the degree of the polynomials in the numerator and denominator are n and k, respectively. Achiezer in 1956 [1].

Theorem

There exists a unique optimal rational function R_{nk}. The number of alternating points is $n + k + 2 - d$, where d is the defect.

$d = \min\{\nu, \mu\}$, $R_{nk} = \frac{P_{n-\nu}}{P_{k-\mu}}$

Rational Approximation methods

- Remez-like method
- Linear inequality method [11]
- Differential correction [3]
- Many other, many of them rely on linear programming.

Optimisation

The objective function is quasiconvex (all the sublevel sets are convex).
Sketch of the proof.

1. R_{nk} is a quasililear (quasiaffine) as a function of the coefficients.
2. The objective function is the supremum for t

$$\max\{f(t) - R_{nk}(A, t), R_{nk}(A, t) - f(t)\}.$$

3. Supremum of quasiconvex functions is quasiconvex.

The optimisation problem is as follows

$$\min_{A,B} \sup_{t\in[c,d]} |f(t) - \frac{A^T G(t)}{B^T H(t)}|, \tag{1}$$

subject to

$$B^T H(t) > 0, \ t \in [c, d], \tag{2}$$

where $f(t) \in C^0_{[c,d]}$ is a function to approximate,
$A = (a_0, a_1, \ldots, a_n)^T \in \mathbb{R}^{n+1}$ and $B = (b_0, b_1, \ldots, b_m)^T \in \mathbb{R}^{m+1}$ are the decision variables,
$G(t) = (g_0(t), \ldots, g_n(t))^T \in \mathbb{R}^{n+1}$ and $H(t) = (h_0(t), \ldots, h_m(t))^T \in \mathbb{R}^{m+1}$ are known functions, in the rest of the paper we refer to them as basis functions. Therefore, we construct the approximations in the form of the ratios of linear combinations of basis functions. Note that the constraint set is an open convex set.

How to reformulate

$$\min z \tag{3}$$

subject to

$$f(t_i) - \frac{A^T G(t_i)}{B^T H(t_i)} \leq z, \quad i = 1, \ldots, N \tag{4}$$

$$\frac{A^T G(t_i)}{B^T H(t_i)} - f(t_i) \leq z, \quad i = 1, \ldots, N \tag{5}$$

$$B^T H(t_i) > 0, \quad i = 1, \ldots, N \tag{6}$$

Now, do the bisection for z, at each iteration z is fixed (Bisection method is coming).

It appeared that Linear Inequality method is Bisection.

Bisection algorithm for quasiconvex optimisation.
Absolute precision for maximal deviation ε.
Set $l \leftarrow 0$
Set u to be maximal deviation for a polynomial approximation (numerator)
$z \leftarrow (u + l)/2$

while $u - l \leq \varepsilon$ **do**
Check feasibility.**if** feasible solution exists **then** $u \leftarrow z$
else
$l \leftarrow z$
end ifupdate $z \leftarrow (u + l)/2$
end whileA, B $\leftarrow$ solve problem with z
return z, A, B

Generalised fractional programming

In generalised fractional programming one needs to minimise supremum of rational functions (subject to linear constraints).

$$f(t_i) - \frac{A^T G(t_i)}{B^T H(t_i)} \leq z, \quad i = 1, \ldots, N \tag{7}$$

$$\frac{A^T G(t_i)}{B^T H(t_i)} - f(t_i) \leq z, \quad i = 1, \ldots, N \tag{8}$$

One of the methods for solving fractional programming problems is Dinkelbach method [7]. There are a number of generalisations of this method to generalised fractional problems and some of them are Differential correction [5].

Bisection vs Differential correction

Bisection method has linear convergence rate (bisecting the maximal deviation at each step).
Differential correction correction is preferable when we are working in the following conditions

1. approximation is univariate;
2. approximation is rational (ratio of two polynomials);
3. full alternation (impossible to know in advance).

There are also a number of specific issues: Differential correction terminates when the improvement (from one iteration to the next one) is small, while Bisection terminates with ε from the optimal solution.

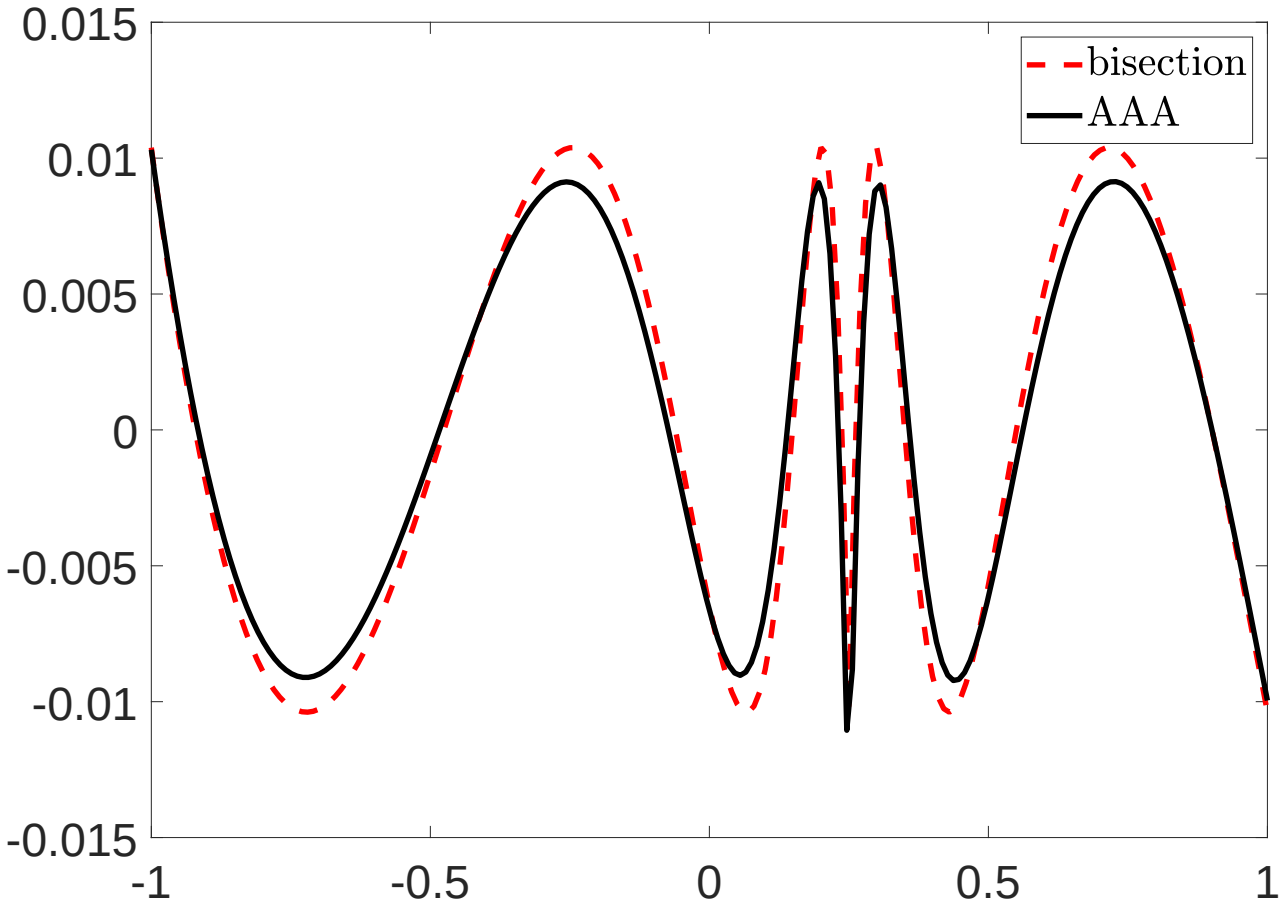
Near optimal approximation

AAA method (Y Nakatsukasa, O. Sete and L. Trefethen).
The name AAA stands for "adaptive Antoulas– Anderson", the scheme is based on [1].

[1] B. Alpert, L. Greengard, and T. Hagstrom, Rapid evaluation of nonreflecting boundary kernels for time-domain wave propagation, SIAM J. Numer. Anal., 37 (2000), pp. 1138–1164.
[2] A. C. Antoulas, Approximation of Large-Scale Dynamical Systems, SIAM, Philadelphia, 2005.
[3] A. C. Antoulas and B. D. Q. Anderson, On the scalar rational interpolation problem, IMAJ. Math. Control Inform., 3 (1986), pp. 61–88.

AAA and Bisection: provided by Vinesha Peiris

Absolute deviation. Approximation function is $f(t) = |t - 0.25|$, $t \in [-1, 1]$



Heat integral

They need to approximate

$$g(m,x) = \int_x^{\infty} \frac{e^{-x}}{x^{m+2}} dx. \tag{9}$$

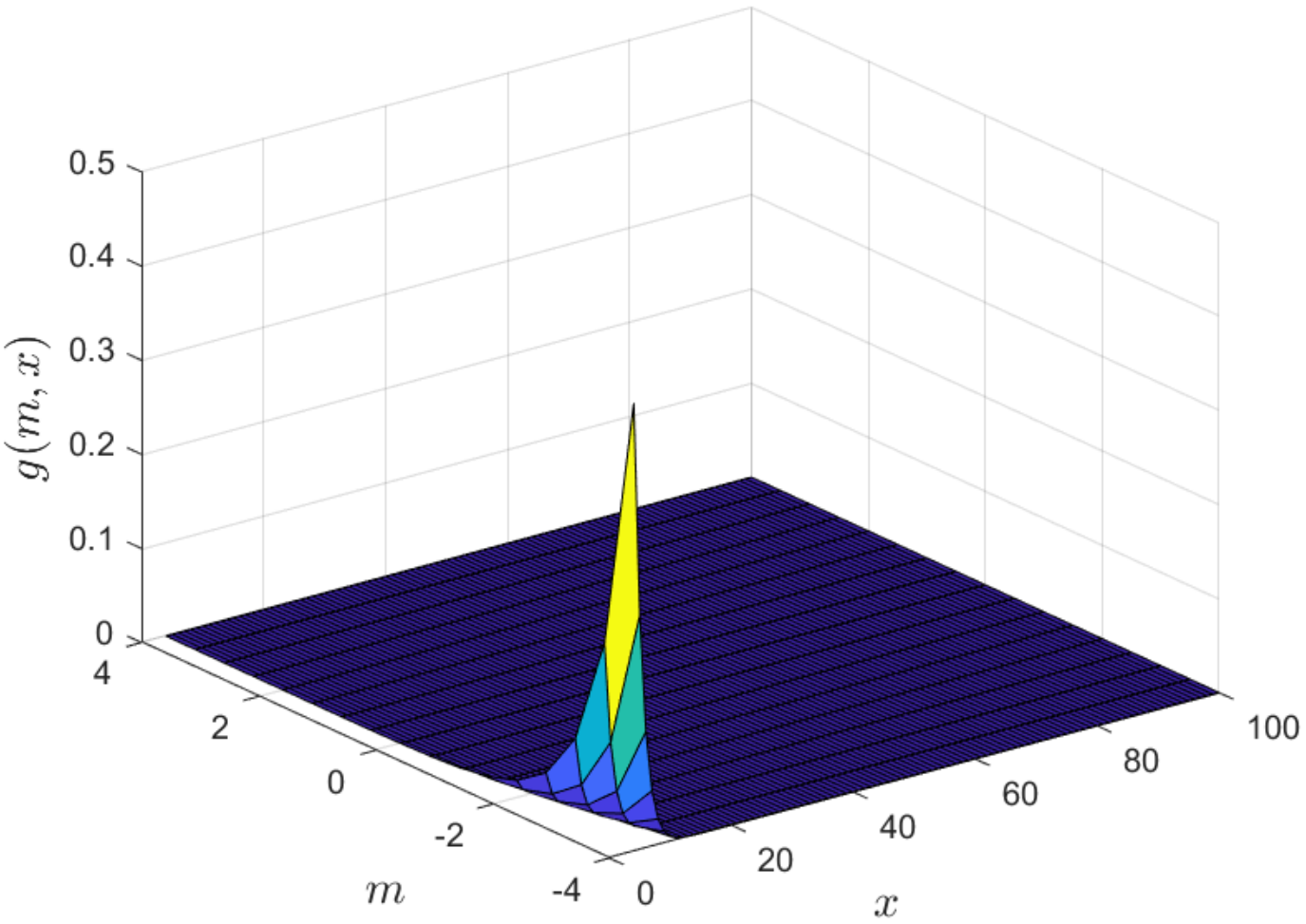
It can be shown that

$$g(m,x) = \frac{e^{-x}}{x^{m+2}} h(m,x). \tag{10}$$

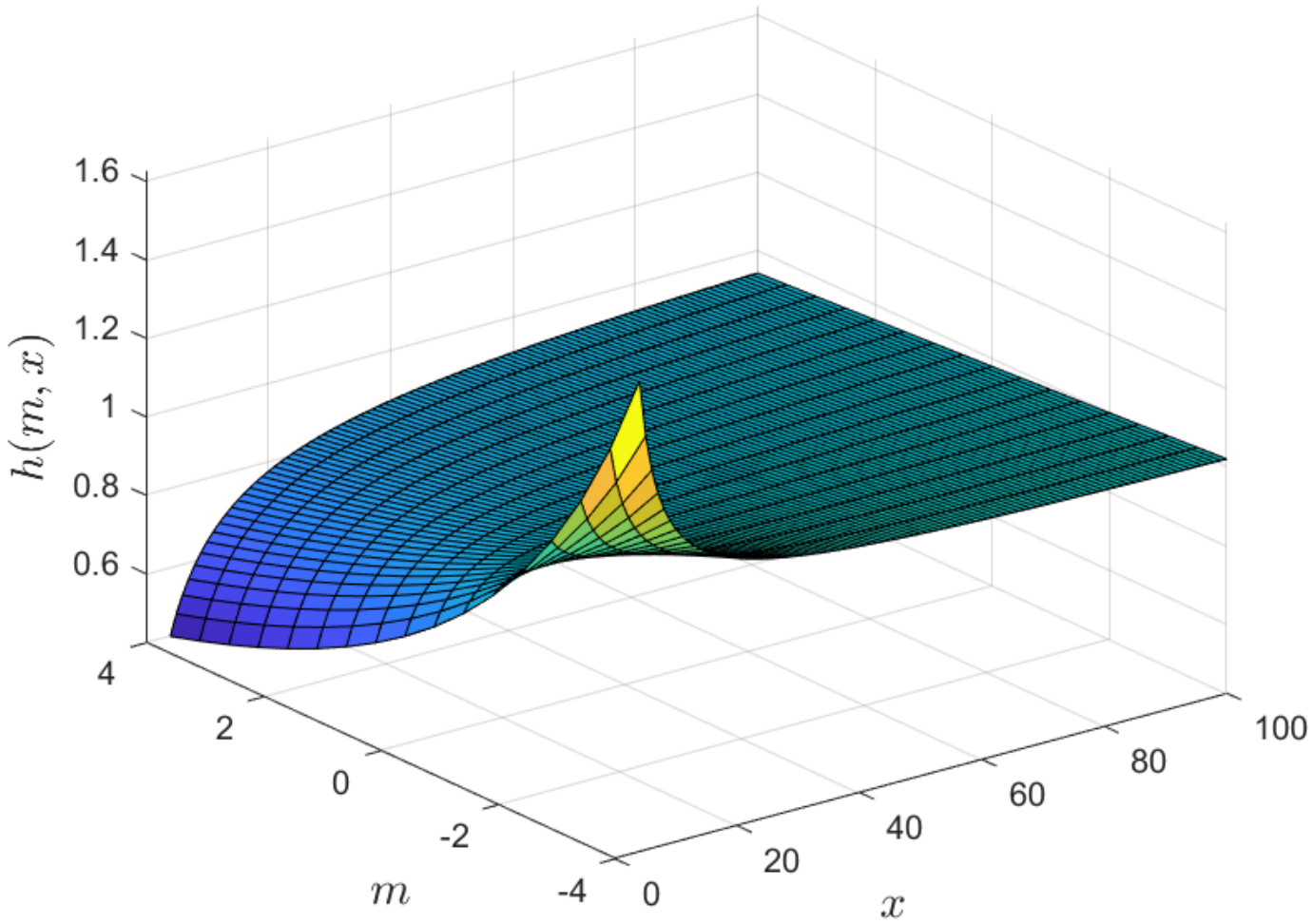
A common approach is to approximate $h(m,x)$ and then multiply by

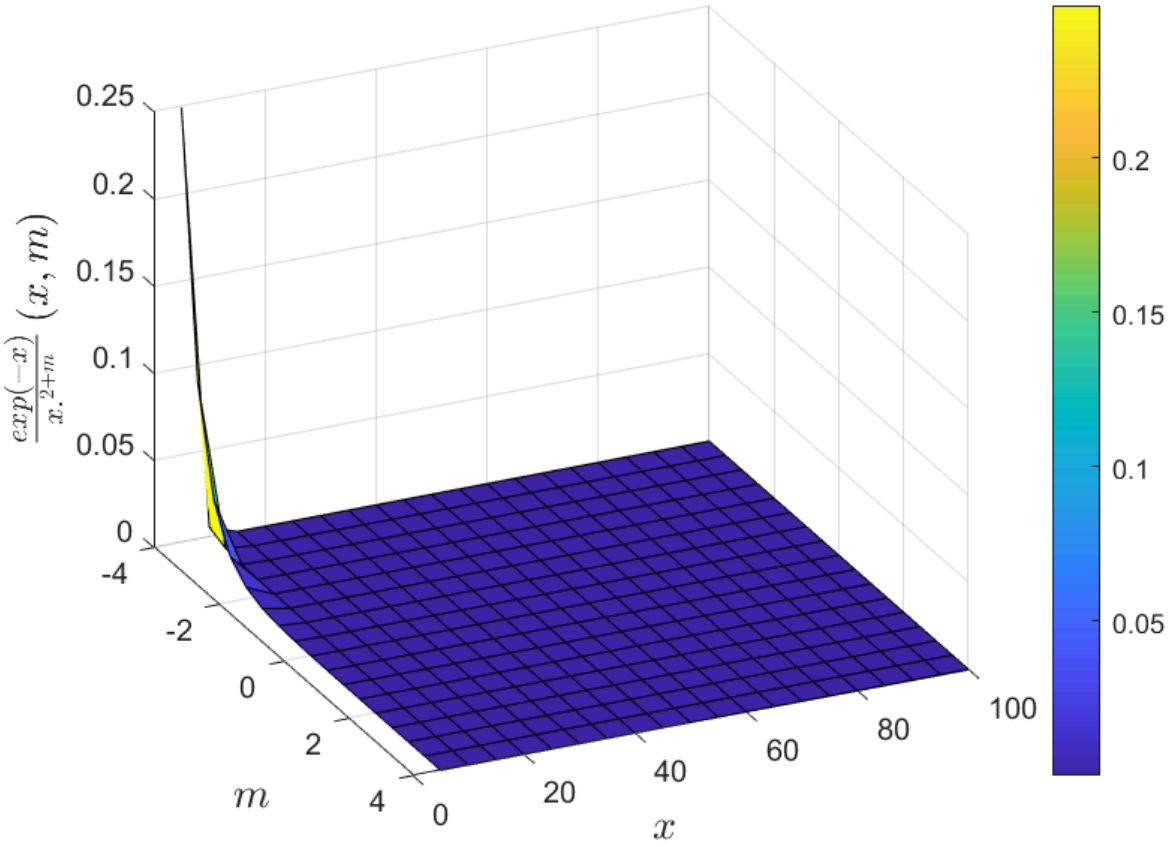
$$\frac{e^{-x}}{x^{m+2}}.$$

$g(m, x)$



$h(m, x)$

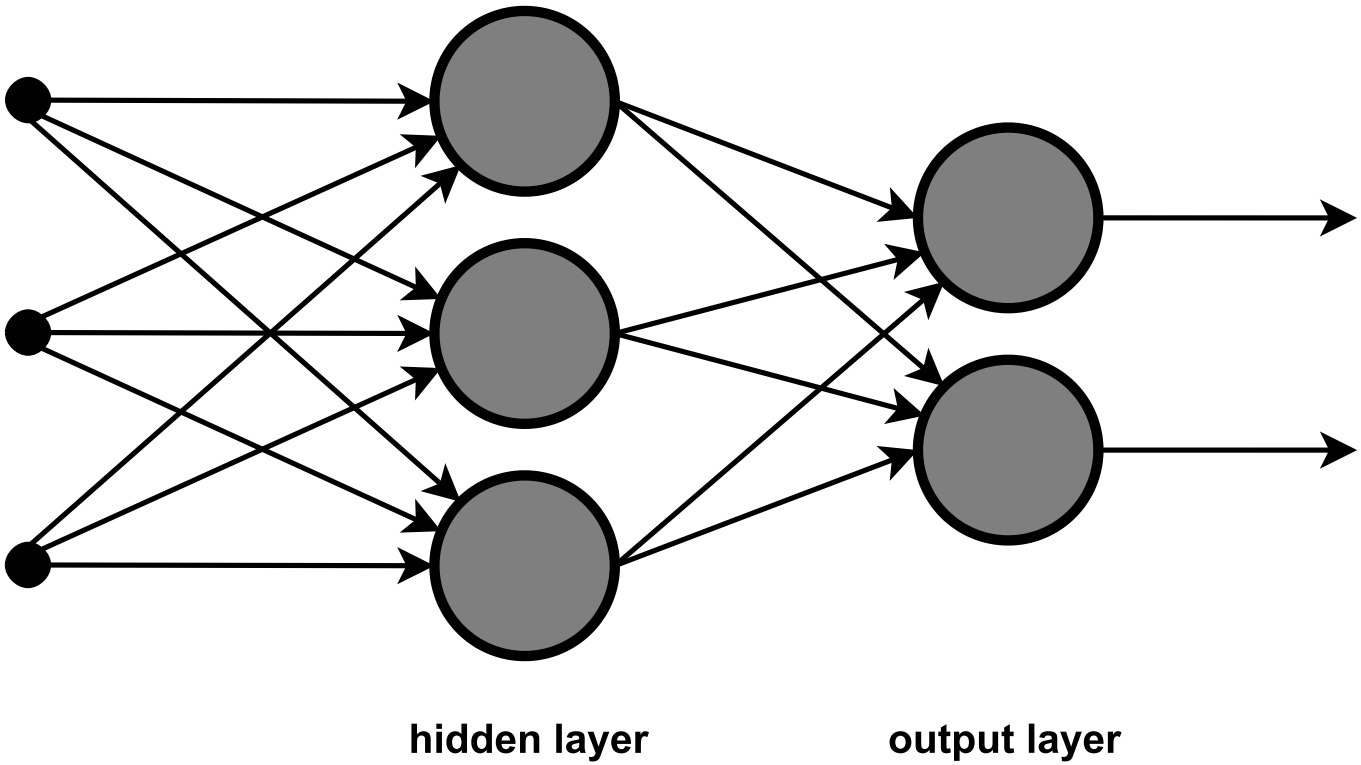



0.25
0.2
0.15
0.1
0.05
0
$\frac{exp(-x)}{x.^{2+m}}(x,m)$
-4
-2
0
2
4
m
0
20
40
60
80
100
x
0.2
0.15
0.1
0.05

Main results

1. Rational approximation of $h(m, x)$ and then multiplication by $\frac{e^{-x}}{x^{m+2}}$
2. Bisection method is more stable than Differential correction.

Neural Network (Wikipedia picture: creative commons)

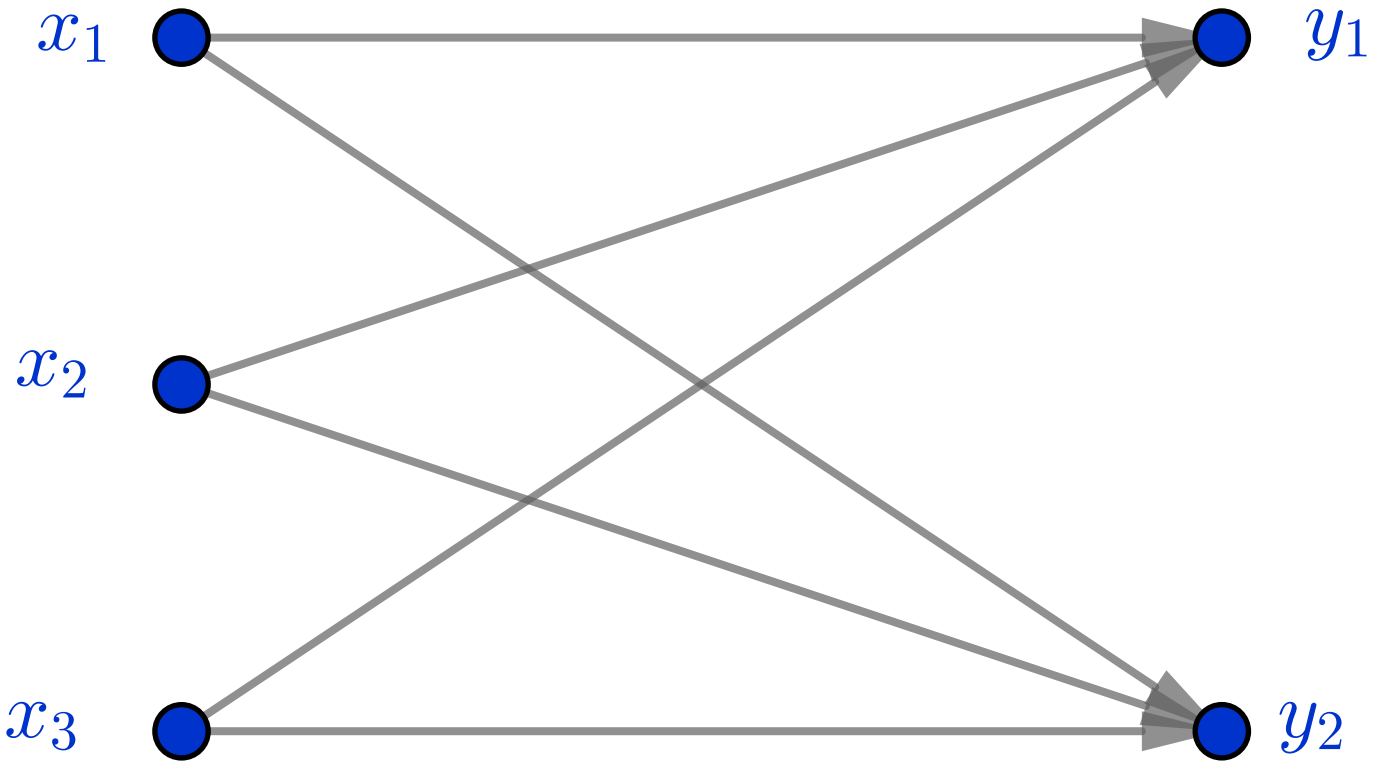


Neural Network: mathematical side

$$\varphi(W, x) = \sigma\left(\sum_{j=1}^{n} w_j x_j^i + w_0\right), \tag{11}$$

- σ is called *activation function*. This is a chosen function, not subject to optimisation.
- W are weights (subject to optimisation).
- If one or more hidden layer is present, φ becomes a composition, where affine transformations are alternating with (different) activation functions.

Neural Network (no hidden layer)



Neural Network and universal approximators: Universal Approximation Theorem

Theorem

Fix a continuous function $\sigma : \mathbb{R} \to \mathbb{R}$ *(activation function) and positive integers* d, D*. The function* σ *is not a polynomial if and only if, for every continuous function* $f : \mathbb{R}^d \to \mathbb{R}^D$ *and every* $\epsilon > 0$ *there exists a continuous function* $f_\epsilon : \mathbb{R}^d \to \mathbb{R}^D$ *(the layer output) with representation* $f_\epsilon = W_2 \circ \sigma \circ W_1$*, where* W_2, W_1 *are affine maps and* $\circ$ *denotes component-wise composition, such that the approximation bound*

$$\sup_{x \in K} \|f(x) - f_\epsilon(x)\| < \varepsilon$$

arbitrarily small (distance from f *to* f_ϵ *can be infinitely small).*

Maths behind Universal approximation

1. A. Kolmogorov and V. Arnold ([9, 2], multivariate function);
2. Cybenko [6], Hoknik [8], Pinkus et al [10] (activation functions);
3. Is Kolmogorov-Arnold theorem relevant to Neural Networks?
4. The most popular ReLU and Leaky ReLU do NOT satisfy the conditions of Cybenko-Hornik-Pinkus theorem (previous slide).

Pinkus et al: Moshe Leshno, Vladimir Ya. Lin, Allan Pinkus, and Shimon Schocken.

Main directions for approximation

1. Approximation of timeseries by rational functions and then use the coefficients of the rational functions as input parameters for the Neural network.
2. Network approximation (no hidden layer) by Leaky ReLU (quasiconvex approximation)
3. Network approximation (no hidden layer) by rational functions.

Approximation of the timeseries: general idea

Generalised rational approximation and its application to improve deep learning classifiers by V Peiris, N Sharon, N Sukhorukova, J Ugon, Applied Mathematics and Computation 389, 125560

$$\min z \tag{12}$$

subject to

$$f(t_i) - \frac{\mathrm{A}^T \mathrm{G}(t_i)}{\mathrm{B}^T \mathrm{H}(t_i)} \leq z, \tag{13}$$

$$\frac{\mathrm{A}^T \mathrm{G}(t_i)}{\mathrm{B}^T \mathrm{H}(t_i)} - f(t_i) \leq z, \tag{14}$$

$$\mathrm{B}^T \mathrm{H}(t_i) > 0, \tag{15}$$

$$t_i \in I, \ \ i = 1, \ldots, N. \tag{16}$$

Note that if (A, B) is an optimal solution, then $\alpha(\mathrm{A}, \mathrm{B}), \ \alpha \in \mathbb{R}$

Approximation of timeseries: one hidden layer

Original (raw data) 4097 features		Model 1 5 features		Model 2 6 features	
Number of nodes	Accuracy	Nodes	Accuracy	Nodes	Accuracy
100 (default)	55%	100	75%	100	75%
1	60%	1	40%	1	40%
2	35%	2	**90%**	2	60%
3	45%	3	60%	3	45%
4	**65%**	4	60%	4	65%
5	45%	5	35%	5	80%
6	45%	6	55%	6	70%
7	55%	7	70%	7	65%
8	55%	8	65%	8	**85%**
9	50%	9	70%	9	65%
10	55%	10	45%	10	45%
2731 (2/3 of inputs)	60%	3	60%	4	65%
8194 (2*inputs+1)	50%	11	75%	13	70%

Approximation of timeseries: two hidden layers

Number of nodes in each layer	Original (raw data) 4097 features	Model 1 5 features	Model 2 6 features
100, 100	65%	60%	45%
2, 1	60%	60%	60%
3, 2	30%	40%	60%
4, 3	60%	45%	75%
5, 1	60%	60%	80%
5, 4	45%	75%	50%
6, 5	70%	40%	80%
7, 6	50%	60%	40%
8, 2	60%	**90%**	40%
8, 7	50%	55%	50%
8, 8	60%	25%	80%
9, 8	45%	55%	80%
10 , 9	45%	75%	**85%**
10, 10	65%	80%	80%
11,11	55%	80%	75%

Approximation of the output using rational functions (no activation function)

This approach is new. Instead of approximating the network by an composition of affine functions and activation functions, we suggest to approximate it by a multivariate rational function.
Kolmogorov-Arnold theorem states that every multivariate continuous function can be written as a finite composition of univariate functions and binary operation of addition:

$$f(\mathrm{x}) = f(x_1, \ldots, x_n) = \sum_{q=0}^{2n} \Phi_q \left(\sum_{p=1}^{n} \phi_{q,p}(x_p) \right).$$

The degree of the monomials does not exceed one.

Approximation of the output using rational functions: experiments

Data	Size	Training	Test	Best result
BirdChicken	512/20/20	100%	55%	98.40%
BeetleFly	512/20/20	100%	80%	94.85%
ElectricDeviceDetect	256/623/3767	100%	85.88%	Unknown
MoteStrain	84/20/1252	100%	75.08%	91.65%
CatsDogs	14773/138/137	70.12%	50.61%	Unknown
Wafer	152/1000/6164	100%	92.98%	99.98%
PowerCons	144/180/180	100%	71.67%	Unknown
ItalyPowerDemand	24/67/1029	100%	93.68%	97.03%
Chinatown	24/20/345	100%	82.80%	Unknown
HandOutlines	2709/1000/370	100%	74.32%	92.37%
DistalPhalanx	80/600/276	63%	58.33%	82.12%
MiddlePhalanx	80/600/291	64.67%	57%	72.23%
SharePriceIncrease	60/965/965	31.30%	31.37%	Unknown

Future research and open problems

- Abstract convexity?
- Approximation by quasiaffine functions (no restriction to rational and generalised rational approximation). The level sets are hyperplanes and the sublevel sets are half-spaces.

The End


Achieser.
Theory of Approximation.
Frederick Ungar, New York, 1956.


Vladimir Arnold.
On functions of three variables.
Dokl. Akad. Nauk SSSR, 114:679–681, 1957.
English translation: Amer. Math. Soc. Transl., 28 (1963), pp. 51–54.


I. Barrodale, M. Powell, and F. Roberts.
The differential correction algorithm for rational l_∞-approximation.
SIAM Journal on Numerical Analysis, 9(3):493–504, 1972.


P. Borwein, I. Daubechies, V. Totik, and G. Nürnberger.
Bivariate segment approximation and free knot splines: Research problems 96-4.
Constructive Approximation, 12(4):555–558, 1996.


J. P. Crouzeix, J. A. Ferland, and S. Schaible.
A note on an algorithm for generalized fractional programs.
Journal of Optimization Theory and Applications, 50:183–187, 1986.


G. Cybenko.
Approximation by superpositions of a sigmoidal function.
Mathematics of Control, Signals and Systems, 2:303–314, 1989.


Werner Dinkelbach.
On nonlinear fractional programming.
Management Science, 13:492–498, 1967.


Kurt Hornik.
Approximation capabilities of multilayer feedforward networks.
Neural Networks, 4(2):251–257, 1991.


A. N. Kolmogorov.
On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition.
Dokl. Akad. Nauk SSSR, 114:953–956, 1957.


Moshe Leshno, Vladimir Ya. Lin, Allan Pinkus, and Shimon Schocken.
Multilayer feedforward networks with a nonpolynomial activation function can approximate any function.
Neural Networks, 6(6):861–867, 1993.


Henry L Loeb.
Algorithms for chebyshev approximations using the ratio of linear forms.
Journal of the Society for Industrial and Applied Mathematics, 8(3):458–465, 1960.


P. Petrushev and V. Popov.
Rational approximation of real functions.
Cambridge University Press, 1987.

Network approximation (quasiconvex model), joint work with Dr. V Roschina and V. Peiris, under review

$$\min_{w\in\mathbb{R}^{n+1}} \max_{i\in 1:N} \left| y^i - \sigma\left(\sum_{j=1}^{n} w_j x_j^i + w_0\right)\right|.$$

σ is the activation function, we work with leaky ReLU, the corresponding optimisation problem is quasiconvex (uniform approximation).

Network approximation (quasiconvex model): numerical experiments (1000 points for training and 370 for the test)

Method	Test set classification accuracy	Confusion matrix
MATLAB toolbox	89.7%	108 | 25 13 | 224
Uniform approximation	60.54%	77 | 56 90 | 147

Table: Original dataset: classification results

Why our procedure is not efficient?

Network approximation (quasiconvex model): comments (1000 points for training and 370 for the test)

Uniform approximation, due to its nature, treats under-represented groups as valid points, while least squares approximation tends to "average" and therefore under-represented groups tend to be "ignored". This is a great advantage when the under-represented groups are outliers, but in many cases these points are valid data.
On the other hand, the presence of ouliers may decrease may decrease the accuracy in the case of uniform approximation. Therefore, our hypothesis is that uniform approximation approach is preferable in the following cases.

1. Absence (or small number) of outliers.
2. Presence of under-represented groups of valid data or uneven distribution of data between the classes (that is, one class is significantly larger than others).
3. Limited size of the available data, where most datapoints are valid and accurate.

Reduced dataset: 5+35 and 35+5

Method	Test accuracy	Confusion matrix	
MATLAB toolbox	64.3%	296	66
		291	347
Uniform approximation	69.5%	193	169
		136	502

Table: Reduced dataset: classification results for uneven number of points from each class in the training set: 5 ponts in Class 1 and 35 points in Class 2.

Method	Test accuracy	Confusion matrix	
MATLAB toolbox	74.6%	116	246
		8	630
Uniform approximation	66.5%	128	234
		101	537

Table: Reduced dataset: classification results for uneven number of points from each class in the training set: 35 points in Class 1 and 5 points in Class 2