Iterative Matching Pursuit and its Applications in Adaptive Time-Frequency Analysis

Zuoqiang Shi

Mathematical Sciences Center, Tsinghua University

Joint work with Prof. Thomas Y. Hou

Outline

- Review of the Data-Driven Time-Frequency Analysis.
- Convergence result of the Data-Driven Time-Frequency Analysis.
- Method for the data with sparse samples and convergence analysis.
- Method for the data with poor scale separation.
- Method for the data with intra-wave frequency modulation and nonlinearity analysis.
- Conclusion remark.
General Framework: sparsest decomposition

**Basic idea:** Looking for the sparsest decomposition over a huge dictionary, $\mathcal{D}$.

Minimize

$$M$$

Subject to:

$$f(t) = \sum_{k=1}^{M} a_k(t) \cos \theta_k(t), \quad a_k(t) \cos \theta_k(t) \in \mathcal{D},$$

(1)
The dictionary is

\[ D = \{ a(t) \cos \theta(t) : \theta'(t) \geq 0, \ a(t) \in V(\theta) \}, \tag{2} \]

where \( V(\theta) \) is a linear space consisting of functions smoother than \( \cos \theta(t) \).

\( V(\theta) \) is chosen to be

\[ V(\theta) = \text{span} \left\{ 1, \cos \left( \frac{k\theta}{2L_\theta} \right), \sin \left( \frac{k\theta}{2L_\theta} \right) : k = 1, \cdots, 2\lambda L_\theta \right\}, \]

where \( \lambda \leq 1/2 \) is a parameter to control the smoothness of functions in \( V(\theta) \) and \( L_\theta = (\theta(T) - \theta(0))/2\pi \).
Algorithm based on \( l^1 \) regularized nonlinear least square

\[ r^0 = f(t). \]

**Step 1:** Solve the following constraint nonlinear least-square problem \((P_2)\):

\[
\text{Minimize} \quad \gamma \| \hat{a}_k \|_1 + \| r^{k-1} - a_k(t) \cos \theta_k(t) \|_2^2 \\
\text{Subject to:} \quad \theta'_k \geq 0, \quad a_k(t) \in V(\theta_k).
\]

where \( \hat{a}_k \) is the representation of \( a_k \) in \( V(\theta_k) \) space

**Step 2:** Update residual

\[
r^k = f(t) - \sum_{j=1}^{k} a_j(t) \cos \theta_j(t)
\]

**Step 3:** If \( \| r^k \|_2 < \epsilon_0 \), stop. Otherwise, go to Step 1.
\[ \theta^0_k = \theta_0. \]

**Step 1:** Get \( a^{n+1}_k \) and \( b^{n+1}_k \) by solving following least square problem:

\[
\min \left\| r^{k-1} - a^{n+1}_k \cos \theta^n(t) - b^{n+1}_k \sin \theta^n(t) \right\|_2^2
+ \gamma\left( \left\| \hat{a}^{n+1}_k \right\|_1 + \left\| \hat{b}^{n+1}_k \right\|_1 \right),
\]

subject to \( a^{n+1}_k(t), b^{n+1}_k(t) \in V(\theta^n_k). \)

where \( \hat{a}^{n+1}_k, \hat{b}^{n+1}_k \) is the coefficients of \( a^{n+1}_k, b^{n+1}_k \) in \( V(\theta^n_k) \) space.

**Step 2:** Update \( \theta^n_k \):

\[
\theta^{n+1}_k = \theta^n_k - \lambda \arctan \left( \frac{b^{n+1}_k}{a^{n+1}_k} \right),
\]

where \( \lambda \in [0, 1] \) is chosen to ensure that \( \theta^{n+1}_k \) is monotonically increasing,

\[
\lambda = \max \left\{ \alpha \in [0, 1] : \frac{d}{dt} \left( \theta^n_k - \alpha \arctan \left( \frac{b^{n+1}_k}{a^{n+1}_k} \right) \right) \geq 0 \right\}.
\]

**Step 3:** If \( \| \theta^{n+1}_k - \theta^n_k \|_2 < \epsilon_0 \), stop. Otherwise, go to Step 1.
Fast algorithm for periodic data

The least square problem can be solved approximately by FFT in $\theta$ space, then we get the following more efficient algorithm:

**Step 1:** Interpolate $r_{k-1}$ from $\{t_i\}_{i=1}^N$ in the physical space to a uniform mesh in the $\theta^n_k$-coordinate to get $r^n_{\theta_k}$ and compute the Fourier transform $\hat{r}^n_{\theta_k}$:

$$r^n_{\theta_k} = \text{Interpolate} \left( r_{k-1}, \theta^n_{k,j} \right),$$

(8)

where $\theta^n_{k,j}, j = 0, \cdots, N - 1$ are uniformly distributed in the $\theta^n_k$-coordinate, i.e. $\theta^n_{k,j} = 2\pi L_{\theta^n_k}^n j/N$. And the Fourier transform of $r^n_{\theta_k}$ is given as follows

$$\hat{r}^n_{\theta_k}(\omega) = \frac{1}{N} \sum_{j=1}^N r^n_{\theta_k,j} e^{-i2\pi \omega \bar{\theta}^n_{k,j}}, \quad \omega = -N/2 + 1, \cdots, N/2,$$

(9)

where $\bar{\theta}^n_{k,j} = \frac{\theta^n_{k,j} - \theta^n_{k,0}}{2\pi L_{\theta^n_k}^n}$.
Fast algorithm for periodic data

**Step 2:** Apply a cutoff function to the Fourier Transform of $r_{\theta_k^n}$ to compute $a$ and $b$ on the mesh of the $\theta_k^n$-coordinate, denoted by $a_{\theta_k^n}$ and $b_{\theta_k^n}$:

$$
a_{\theta_k^n} = \mathcal{F}^{-1} \left[ (\hat{r}_{\theta_k^n} (\omega + L_{\theta_k^n}) + \hat{r}_{\theta_k^n} (\omega - L_{\theta_k^n})) \cdot \chi_\lambda (\omega / L_{\theta_k^n}) \right],$$

$$
b_{\theta_k^n} = \mathcal{F}^{-1} \left[ i \cdot (\hat{r}_{\theta_k^n} (\omega + L_{\theta_k^n}) - \hat{r}_{\theta_k^n} (\omega - L_{\theta_k^n})) \cdot \chi_\lambda (\omega / L_{\theta_k^n}) \right].$$

$\mathcal{F}^{-1}$ is the inverse Fourier transform defined in the $\theta_k^n$ coordinate:

$$
\mathcal{F}^{-1} (\hat{r}_{\theta_k^n}) = \frac{1}{N} \sum_{\omega=-N/2+1}^{N/2} \hat{r}_{\theta_k^n} e^{i2\pi\omega\theta_k^n} j, \quad j = 0, \cdots, N - 1.
$$

**Step 3:** Interpolate $a_{\theta_k^n}$ and $b_{\theta_k^n}$ from the uniform mesh $\{\theta_k^n, j\}_{j=1}^N$ in the $\theta_k^n$-coordinate back to the physical grid points $\{t_i\}_{i=1}^N$:

$$
a(t_i) = \text{Interpolate} \left( a_{\theta_k^n}, t_i \right), \quad i = 0, \cdots, N - 1,$$

$$
b(t_i) = \text{Interpolate} \left( b_{\theta_k^n}, t_i \right), \quad i = 0, \cdots, N - 1.$$
Fast algorithm for periodic data

Step 4: Update $\theta^n$ in the $t$-coordinate:

$$\Delta \theta' = P_{V(\theta; \eta)} \left( \frac{d}{dt} \left( \arctan \left( \frac{b_k^{n+1}}{a_k^{n+1}} \right) \right) \right), \quad \Delta \theta(t) = \int_0^t \Delta \theta'(s) ds,$$

$$\theta_k^{n+1} = \theta_k^n - \beta \Delta \theta,$$

where $\beta \in [0, 1]$ is chosen to make sure that $\theta_k^{n+1}$ is monotonically increasing:

$$\beta = \max \left\{ \alpha \in [0, 1] : \frac{d}{dt} \left( \theta_k^n - \alpha \Delta \theta \right) \geq 0 \right\}.$$

and $P_{V_p(\theta; \eta)}$ is the projection operator to the space $V_p(\theta; \eta)$.

Step 5: If $\|\theta_k^{n+1} - \theta_k^n\|_2 < \epsilon_0$, go to step 6. Otherwise, set $n = n + 1$ and go to Step 1.

Step 6: If $\eta \geq \lambda$, stop. Otherwise, set $\eta = \eta + \Delta \eta$ and go to step 1.
Convergence analysis

Theorem (Exact recovery)

For the data $f = f_0 + f_1 \cos \theta + \sigma(t)$, $|\sigma(t)| \leq \epsilon$ which is noise. Assume that the instantaneous frequency $\theta'$ is $M_0$-sparse over the Fourier basis in physical space, $\theta' \in V_{M_0} = \text{span}\{e^{i2\pi k T}, k = -M_0, \cdots, 1, \cdots, M_0\}$, mean $f_0$ and envelop $f_1$ are $M_1$-sparse over the Fourier basis in $\theta$-space,

$$\hat{f}_{0,\theta}(k) = \hat{f}_{1,\theta}(k) = 0, \quad \forall |k| > M_1.$$

If the initial guess of $\theta$ satisfies

$$\|F\left((\theta^0 - \theta)\right)\|_1 \leq \pi M_0 / 2,$$

then there exist $\eta_0 > 0$ such that

$$\|F\left((\theta^{m+1} - \theta)\right)\|_1 \leq \frac{1}{2} \|F\left((\theta^m - \theta)\right)\|_1 + C \cdot \epsilon,$$

provided $L = \frac{\theta(T) - \theta(0)}{2\pi} \geq \eta_0(M_0, M_1, \epsilon)$, where $\eta_0(M_0, M_1, \epsilon)$ is a constant determined by $M_0$, $M_1$, $\epsilon$. 

Zuoqiang Shi, MSC, Tsinghua

Adaptive Data Analysis
Numerical Validation

\[ \theta = 20\pi t + 2 \cos 2\pi t + 2 \sin 4\pi t, \quad \bar{\theta} = \theta/10 \]

\[ f_0 = 2 + \cos \bar{\theta} + 2 \sin 2\bar{\theta} + \cos 3\bar{\theta}, \quad f_1 = 3 + \cos \bar{\theta} + \sin 3\bar{\theta} \]

\[ f = f_0 + f_1 \cos \theta. \]

**Figure:** Original data and errors of IMF and phase function.
Theorem (Approximate recovery)

Assume that the Fourier coefficients of the instantaneous frequency $\theta'$, local mean $f_0$ and envelop $f_1$ all have fast decay, i.e. there exists $C_0 > 0$, $p \geq 4$ such that

$$|\mathcal{F}(\theta')(k)| \leq C_0|k|^{-p}, \quad |\mathcal{F}_\theta(f_0)(k)| \leq C_0|k|^{-p}, \quad |\mathcal{F}_\theta(f_1)(k)| \leq C_0|k|^{-p}$$

If $L$ is large enough and the initial guess satisfies

$$\|\mathcal{F}((\theta^0 - \theta)')\|_1 \leq \pi M_0/2,$$

then, we have

$$\|\mathcal{F}((\theta^{m+1} - \theta)')\|_1 \leq \Gamma_0(L/4)^{-p+2} + C_0 M_0^{-p+1} + \frac{1}{2} \|\mathcal{F}((\theta^m - \theta)')\|_1,$$

where $\Gamma_0 > 0$ is a constant determined by $C_0$, $M_0$ and $f_1$. 

Zuoqiang Shi, MSC, Tsinghua
Adaptive Data Analysis
Numerical Validation

\[ a_0 = \frac{1}{1.1 + \sin(2\pi t)}, \quad a_1 = \frac{1}{1.1 + \cos(2\pi t)}, \]

\[ \theta = 10\sin(2\pi t) + 40\pi t, \quad f = a_0 + a_1 \cos \theta \]

**Figure:** IMF and instantaneous frequency with bad scale separation.
Numerical Validation

\[ f(t) = a_0(t) + a_1(t) \cos(2\theta(t)), \] where \( a_0(t), a_1(t) \) and \( \theta(t) \) are the same as the previous example.

**Figure:** IMF and instantaneous frequency with good scale separation.
Another Formulation: Basis Pursuit

From another point of view, the optimization problem can be seen as the nonlinear version of the $L_0$ minimization problem:

$$\min_{x,\theta_1,\cdots,\theta_M} \|x\|_0, \quad \text{subject to} \quad [A_{\theta_1}, \cdots, A_{\theta_M}] x = f.$$ 

where $A_{\theta_j}$ is the collection of the (overcomplete) Fourier basis in $\theta_j$ coordinate.

One natural idea to solve above optimization problem is to solve a $l^1$ optimization problem with fixed $\theta_1, \cdots, \theta_M$ and update $\theta_1, \cdots, \theta_M$ iteratively.

- Solve

$$\min_x \|x\|_1, \quad \text{subject to} \quad [A_{\theta_1^n}, \cdots, A_{\theta_M^n}] x = f.$$ 

- Update $\theta_1^n, \cdots, \theta_M^n$ to get $\theta_1^{n+1}, \cdots, \theta_M^{n+1}$. 

Zuoqiang Shi, MSC, Tsinghua

Adaptive Data Analysis
Another Formulation: Basis Pursuit

Corresponding to the algorithm based on the matching pursuit, the phase function is calculated one by one to reduce the complexity of the method,

$$
\min_{x, \theta} \|x\|_0, \quad \text{subject to} \quad A_{\theta}x = f.
$$

where $A_{\theta}$ is the collection of the (overcomplete) Fourier basis in $\theta$ coordinate.

The algorithm should be

- Solve

$$
\min_x \|x\|_1, \quad \text{subject to} \quad A_{\theta^n}x = f.
$$

- Update $\theta^n$ to get $\theta^{n+1}$.

This gives us the immediate generalization for the data with sparse samples.
To deal with the data with sparse samples, we only need to replace the interpolation-FFT by solving a $l^1$ minimization problem.

Solve the $l_1$ minimization problem to get the Fourier transform of the signal $f$ in $\theta^m$ coordinate:

$$\hat{f}_{\theta^m} = \arg \min_{x \in \mathbb{R}^{N_b}} \|x\|_1, \quad \text{subject to} \quad A_{\theta^m} \cdot x = f$$

where $A_{\theta^m} \in \mathbb{R}^{N_s \times N_b}$, $N_s < N_b$, $N_s$ is the number of samples and $N_b$ is the number of Fourier basis.

$$A_{\theta^m}(j, k) = e^{i2\pi k \bar{\theta}^m(t_j)}, \quad j = 1, \cdots, N_s, \quad k = -N_b/2 + 1, \cdots, N_b/2$$

and $\bar{\theta}^m = \frac{\theta^m(\bar{T}) - \theta^m(0)}{\theta^m(T) - \theta^m(0)}$. 
\[
\theta(t_i) = 120\pi t_i + 10 \cos(2\pi t_i), \quad a(t_i) = 2 + \cos(2\pi t_i), \quad f(t_i) = a(t_i) \cos \theta(t_i),
\]

and \(i = 1, 2, \cdots, N\).

In this example, the number of samples is 64, approximately one sample point in one period of the signal. The location \(t_i\) is chosen randomly in \([0, 1]\).
Figure: Left: original samples, red: exact; blue: recovered; '*' represent the sample points. Right: instantaneous frequency, red: exact; blue: numerical.
Data with sparse samples

Figure: Left: original samples, $f(t) + 0.2X(t)$, red: exact; blue: recovered from the noised data; '*' represent the sample points. Right: instantaneous frequency, red: exact; blue: numerical.
Data with sparse samples

Theorem

If the sample points are selected at random, under the same assumption in Theorem 1, we have there exist $\eta_0 > 0$, $\eta_1 > 0$, such that with overwhelming probability

$$\| F \left( (\theta^{m+1} - \theta)' \right) \|_1 \leq \frac{1}{2} \| F \left( (\theta^m - \theta)' \right) \|_1,$$

(11)

provided $L \geq \eta_0$ and $N_s \geq \eta_1 \max(\bar{\theta})' (\log N_b)^6$, $N_s$ is the number of samples, $N_b$ is the number of basis.
Using the assumption that $a(t)$ is smooth, we put a weight in the $l^1$ term to penalize the high wave number components,

- Solve

$$\min_x \| \omega x \|_1, \quad \text{subject to } A_{\theta^n} x = f.$$ 

where $\omega$ is a weight vector.

- Update $\theta^n$ to get $\theta^{n+1}$.

We can get better theoretical estimate for this formulation.
Data with sparse samples

Theorem

If the samples points are selected at random, under the same assumption in Theorem 1, we have there exist $\eta_0 > 0$, $\eta_1 > 0$, such that with probability more than $1 - \delta$

$$\left\| \mathcal{F} \left( (\theta^{m+1} - \theta)' \right) \right\|_1 \leq \frac{1}{2} \left\| \mathcal{F} \left( (\theta^m - \theta)' \right) \right\|_1,$$

(12)

provided $L \geq \eta_0$ and $N_s \geq \eta_1 \max(\bar{\theta})' \max \{ \log N_b, - \log \delta \}$, $N_s$ is the number of samples, $N_b$ is the number of basis.
Data with incomplete samples

\[ \theta(t) = 120 \pi t + 10 \cos(4\pi t), \quad a(t) = 2 + \cos(2\pi t), \]
\[ f(t) = a(t) \cos \theta(t), \quad t \in [0, 0.4] \cup [0.6, 1]. \]

**Figure:** Left: blue: original incomplete data, the gap is (0.4, 0.6); red: missing data obtained by our method; Middle: recovered missing data, red: exact; blue: numerical. Right: recovered instantaneous frequency, red: exact; blue: numerical.
Data with incomplete samples

Figure: Left: blue: the original incomplete data, the gap is (0.3, 0.7); red: the missing data recovered by our method; Middle: the recovered missing data, red: exact; blue: numerical. Right: the instantaneous frequencies, red: exact; blue: numerical.
Data without scale-separation

\[ f(t) = \cos(20\pi t + 40\pi t^2 + \sin(2\pi t)) + \cos(40\pi t) \]  

(13)

**Figure:** Original data.
Figure: Left: Instantaneous frequencies; red: exact frequencies; blue: numerical results. Middle and Right: IMFs extracted by previous method.
Data without scale-separation

For this kind of signals, we have to decompose these components simultaneously, since they have strong correlation.

- Solve
  \[
  \min_x \|x\|_1, \quad \text{subject to} \quad [A_{\theta_1^n}, \ldots, A_{\theta_M^n}] x = f.
  \]

- Update \(\theta_1^n, \ldots, \theta_M^n\) to get \(\theta_1^{n+1}, \ldots, \theta_M^{n+1}\).

Combining the method to update the phase function, we can get following algorithm.
Initialize: $n = 0, \quad \eta = 0$.

**Step 1:** Solve the following $l^1$ regularized least-square problem:

$$(a_k^{n+1}, b_k^{n+1}) \in \operatorname{Argmin}_{a_k, b_k} \gamma \sum_{k=1}^{M} \left( \| \hat{a}_k \|_1 + \| \hat{b}_k \|_1 \right) + \| f - \sum_{k=1}^{M} (a_k \cos \theta^n_k + b_k \sin \theta^n_k) \|_2^2$$

Subject to: $a_k \in V(\theta^n_k), \quad b \in V(\theta^n_k)$,

where $\hat{a}_k, \hat{b}_k$ are the representations of $a_k, b_k$ in the $V(\theta^n_k)$ space.
Step 2: Update $\theta^n_k$:

$$\Delta \theta'_k = P_V(\theta^n_k; \eta) \left( \frac{d}{dt} \left( \arctan \left( \frac{b_k^{n+1}}{a_k^{n+1}} \right) \right) \right),$$  \hfill (14)  

$$\Delta \theta_k = \int_0^t \Delta \theta'_k(s) \, ds, \quad \theta^{n+1}_k = \theta^n_k - \beta_k \Delta \theta_k,$$  \hfill (15)

where $\beta_k \in [0, 1]$ is chosen to make sure that $\theta^{n+1}_k$ is monotonically increasing:

$$\beta_k = \max \left\{ \alpha \in [0, 1] : \frac{d}{dt} (\theta^n_k - \alpha \Delta \theta_k) \geq 0 \right\}. \hfill (16)$$

and $P_V(\theta^n_k; \eta)$ is the projection operator to the space $V(\theta^n_k; \eta)$.

Step 3: If $\sum_{k=1}^M \| \theta^{n+1}_k - \theta^n_k \|_2 > \epsilon_0$, set $n = n + 1$ and go to Step 1. Otherwise, go to step 4.

Step 4: If $\eta \geq \lambda$, stop. Otherwise, set $\eta = \eta + \Delta \eta$ and go to step 1.
Data without scale-separation

Figure: Left: Instantaneous frequencies; Middle and right: IMFs extracted by extracting two IMFs together. red: exact results; blue: numerical results.
In the previous approach, we want to decompose the signal to the form

\[ f(t) = \sum_{k=1}^{M} a_k(t) \cos \theta_k(t). \]  

(17)

For some applications, decompositions of above form are too restrictive. For example, the decomposition of following ECG data is not sparse at all.

**Figure:** Typical ECG data
Using the concept of the shape function introduced by Prof. Daubechies and Wu, we decompose the signal to the following form

\[
f(t) = \sum_{k=1}^{M} a_k(t) s_k(\theta_k(t)).
\]  

(18)

where \(s_k(\cdot)\) are \(2\pi\)-period function.

\(a_k, \theta_k, s_k\) can be obtained by solving following optimization problem:

\[
\min_{a_k, \theta_k, s_k} \| f(t) - \sum_{k=1}^{M} a_k(t) s_k(\theta_k(t)) \|^2_2,
\]

subject to: \(a_k, \theta_k\) are smoother than \(\cos \theta_k\),

\(s_k(\cdot)\) is \(2\pi\)-periodic.
First, we consider a simple case, let $M = 1$, then we only need to solve

$$\min_{a, \theta, s} \| f(t) - a(t)s(\theta(t)) \|_2^2,$$

subject to: $a, \theta \in V(\theta)$, $s(\cdot)$ is $2\pi$-periodic.

The phase function $\theta$ can be obtained by the method introduce before, then

$$\min_{a, s} \| f(t) - a(t)s(\theta(t)) \|_2^2,$$

subject to: $a \in V(\theta)$, $s(\cdot)$ is $2\pi$-periodic.
Since $s$ is periodic, it can be represented by Fourier basis,

$$s(\theta) = \sum_{k=-K}^{K} c_k e^{ik\theta}$$

(19)

Using this representation, the optimization problem becomes

$$\min_{a, c_k} \| f - a \sum_{k=-K}^{K} c_k e^{ik\theta} \|_{2,\theta}^2,$$

subject to: $a \in V(\theta)$. 

Zuoqiang Shi, MSC, Tsinghua
Adaptive Data Analysis
Using the Parseval equality,

\[
\min_{\hat{a}, c_k} \left\| \hat{f}_\theta(\omega) - \sum_{k=-K}^{K} c_k \hat{a}(\omega + kL_\theta) \right\|^2_2,
\]

subject to: \( a \in V(\theta) \).

where \( L_\theta = (\theta(T) - \theta(0))/2\pi \).

Using the constraint that \( \hat{a}(\omega) = 0, |\omega| > \lambda L_\theta \) and \( \lambda \leq 1/2 \).

\[
\min_{\hat{a}, c_k} \sum_{k=-K}^{K} \sum_{|\omega|<\lambda L_\theta} |\hat{f}_\theta(\omega + kL_\theta) - c_k \hat{a}(\omega)|^2,
\]

which can be solved by SVD.
Algorithm to compute $s$ and the envelop $a$,

- Compute the phase function $\theta$.

- Interpolate the signal $f$ to $\theta$ coordinate and apply FFT to get the Fourier coefficients of $f$ over $\theta$ coordinate, $\hat{f}_\theta$.

- Chop $\hat{f}_\theta$ to several pieces to form a matrix $\hat{F}_\theta$.

- Apply the singular value decomposition on $\hat{F}_\theta$ to get $c_k$ and $\hat{a}$. 
Data with intra-wave frequency modulation: Duffing equation

Figure: Left: The solution of the Duffing equation; Right: The shape function obtain by our method.
Data with intra-wave frequency modulation: Duffing equation

**Figure:** Left: The solution of the Duffing equation plus noise; Right: The shape function obtain by our method.
Data with intra-wave frequency modulation: ECG data

Figure: Left: The original ECG data; Right: The shape function obtained by our method for the ECG data.
For each IMF, there exists a second order ODE,

$$\ddot{x} + a(x, t)\dot{x} + b(x, t) = 0$$  \hspace{1cm} (20)

and $a(x, t)$, $b(x, t)$ are slowly varying with respect to $t$.

If we looking for $a(x, t)$ and $b(x, t)$ locally, they can be approximate by two functions which are independent on $t$, then the ODE becomes autonomous,

$$\ddot{x} + a(x)\dot{x} + b(x) = 0.$$  \hspace{1cm} (21)

or we can rewrite it to the conservative form,

$$\ddot{x} + \dot{Q}(x) + b(x) = 0.$$  \hspace{1cm} (22)
Then we can get weak formulation of the ODE,

\[ \langle x, \ddot{\phi} \rangle - \langle Q(x), \dot{\phi} \rangle + \langle b(x), \phi \rangle = 0, \quad \forall \phi \in C_0^\infty([0, T]). \]  

(23)

In order to determine the ODE, we use polynomials to approximate \( Q(x) \) and \( b(x) \),

\[ Q(x) = \sum_{k=0}^{M} q_k x^{k+1}, \quad b(x) = \sum_{k=0}^{M} b_k x^k \]  

(24)

where \( M \) is the order of polynomials which is given \( a \ priori \), \( q_k \), \( b_k \) are constants.
Using this formulation, we can design following optimization problem to solve $a_k$ and $b_k$, 

$$(q_k, b_k) = \arg \min_{\alpha_k, \beta_k} \gamma \sum_{k=1}^{M} (|\alpha_k| + |\beta_k|)$$

$$+ \sum_{i=1}^{N} \left| \langle x, \phi_i \rangle - \sum_{k=0}^{M} \alpha_k < x^{k+1}, \phi_i > + \sum_{k=0}^{M} \beta_k < x^k, \phi_i > \right|^2$$

The test function we use is the following cosine type function, 

$$\phi_i(t) = \begin{cases} 
\frac{1}{2}(1 + \cos(\pi(t - t_i)/\lambda)), & -\lambda < t - t_i < \lambda, \\
0, & \text{otherwise.}
\end{cases}$$

where $t_i, i = 1, \cdots, N$ is the centers of the test functions and the parameter $\lambda$ determines their support.
The Duffing equation is a nonlinear ODE which has the following form:
\[
\ddot{u} + u + u^3 = 0
\]  

(25)
\[u(0) = 1, \dot{u}(0) = 0.\]

**Figure:** Top: the solution of the Duffing equation; Middle: Coefficients of polynomials; Bottom: Degrees of nonlinearity.
The Van der Pol equation is a nonlinear ODE which has the following form:

\[ \ddot{u} + (u^2 - 1)\dot{u} + u = 0 \]  

(26)

\[ u(0) = 1, \ u'(0) = 0. \]

**Figure**: Top: the solution of the Van der Pol equation; Middle: Coefficients of polynomials; Bottom: Degrees of nonlinearity.
Consider a ODE with varying coefficients,

\[ \ddot{u} + a(t)(u^2 - 1)\dot{u} + (1 - a(t))u^3 + u = 0 \]  

(27)

where \( a(t) = \frac{1}{2} \left( 1 - \frac{t-100}{\sqrt{(t-100)^2+400}} \right) \). The initial condition is that \( \dot{u}(0) = 0, u(0) = 1 \) and the equation is solved over \( t \in [0, 200] \).

![Figure: Zuoqiang Shi, MSC, Tsinghua Adaptive Data Analysis](image-url)
Consider a ODE with varying coefficients,

\[
\ddot{u} + a(t)(u^2 - 1)\dot{u} + (1 - a(t))u^3 + u = 0 \tag{28}
\]

where \(a(t) = (1 - \text{sgn}(t - 100))/2\). The initial condition is that \(\dot{u}(0) = 0, u(0) = 1\) and the equation is solved over \(t \in [0, 200]\).
Data with intra-wave frequency modulation: Nonlinear Dynamical System

Figure:

Zuoqiang Shi, MSC, Tsinghua
Adaptive Data Analysis
Concluding Remarks

- We generalize the data-driven time-frequency analysis method proposed by Hou and Shi to several more complicated data sets.

- Convergence analysis has been performed under some scale separation assumption on the multiscale data.

- One nonlinearity analysis method based on second order ODE has been developed to deal with the nonlinear data.

- Applications to biomedical problems, climate data, and geophysical problems are under investigation.
Reference:


- T. Y. Hou and Z. Shi, Data-driven Time-Frequency Analysis, to appear in ACHA.

Thank you!