Supervised learning of a regression model based on latent process. Application to the estimation of Fuel Cell lifetime

Raïssa Onanena(1), Faïcel Chamroukhi(1), Latifa Oukhellou(1), Denis Candusso(1,3), Patrice Aknin(1), Daniel Hissel(2,3)

INRETS-LTN, 2 av de la butte verte, 93166 Noisy le Grand Cedex, France
FEMTO-ST UMR CNRS 6174, Université de Franche-Comté, 90010 Belfort, France
FCLAB, Rue Ernest Thierry-Mieg, 90010 Belfort Cedex, France

December 15 2009
1 Introduction
   • Context
   • Available data

2 Feature extraction
   • A probabilistic approach
   • Parameter estimation

3 Fuel Cell lifetime estimation

4 Conclusion
Context: Predictive maintenance of the Fuel Cells (FCs)

- Fuel Cells (FCs) are widely used in many domains including transport
- They can offer high fuel economy
- Lower CO$_2$ emissions
- The stack is affected by the operating conditions (temperature, mechanical constraints on the membrane, electrode assemblies etc.)

⇒ a predictive maintenance policy is needed

Aim

FC lifetime estimation using specific measurements acquired during the ageing study of the stacks
The Electrochemical impedance spectrum (EIS)

- Measurements of the Electrochemical Impedance Spectrum (EIS) are generally used for FC characterization

The impedance spectrum for the Fuel Cell consists of three regimes:
- a first capacitive arc ($f < 130$Hz) due to the diffusion phenomena
- a second capacitive arc ($130$Hz $\leq f < 4$kHz) linked to the FC membrane charges
- a last inductive part arc which is present in high frequencies ($4$kHz $\leq f$) due to the inductive behavior of connections
Evolution of the real and imaginary parts of the Electrochemical Impedance Spectrum (EIS) over time

Real part

Imaginary part
Feature extraction from the imaginary part of the EIS

- The imaginary part of the spectrum is more informative and more complex than the real part.
- Particularly, three regimes corresponding to the behaviour of the stack are perceptible:

![Graph showing frequency vs. Im(Z)](image)

- Smooth or abrupt changes between the different regimes.
Feature extraction from the imaginary part of the EIS

- The imaginary part of the spectrum is more informative and more complex than the real part.
- Particularly, three regimes corresponding to the behaviour of the stack are perceptible:

![Graph showing the imaginary part of the spectrum against frequency.]

- Smooth or abrupt changes between the different regimes.

⇒ The proposed solution: use an adapted regression model whose parameters will be used as the feature vector for each EIS.
A regression model with a hidden logistic process

The data: \( \{(x_1, f_1), \ldots, (x_n, f_n)\} \)

- \( x_i \): real dependent variable: - Imaginary part of the EIS
- \( f_i \): independent variable representing the frequency

\[
\forall i = 1, \ldots, n, \quad x_i = \beta_{z_i}^T r_i + \sigma_{z_i} \epsilon_i \quad ; \quad \epsilon_i \sim \mathcal{N}(0, 1),
\]

- \( z_i \in \{1, \ldots, K\} \) hidden variable: the class label of the component generating \( x_i \)
- \( \beta_{z_i} \in \mathbb{R}^{p+1} \): regression coefficients of the sub-model \( z_i \)
- \( r_i = (1, f_i, \ldots, f_i^p)^T \): covariate vector in \( \mathbb{R}^{p+1} \)
A regression model with a hidden logistic process

The data: \( \{(x_1, f_1), \ldots, (x_n, f_n)\} \)

- \( x_i \): real dependent variable: - Imaginary part of the EIS
- \( f_i \): independent variable representing the frequency

\[ \forall i = 1, \ldots, n, \quad x_i = \beta_{z_i}^T r_i + \sigma_{z_i} \epsilon_i ; \quad \epsilon_i \sim \mathcal{N}(0, 1), \]

- \( z_i \in \{1, \ldots, K\} \) hidden variable: the class label of the component generating \( x_i \)
- \( \beta_{z_i} \in \mathbb{R}^{p+1} \): regression coefficients of the sub-model \( z_i \)
- \( r_i = (1, f_1, \ldots, f_i^p)^T \): covariate vector in \( \mathbb{R}^{p+1} \)

\( z = (z_1, \ldots, z_n) \) is a hidden logistic process
A regression model with a hidden logistic process

The data: \( \{(x_1, f_1), \ldots, (x_n, f_n)\} \)

- \( x_i \): real dependent variable: - Imaginary part of the EIS
- \( f_i \): independent variable representing the frequency

\[
\forall i = 1, \ldots, n, \quad x_i = \beta_{z_i}^T r_i + \sigma_{z_i} \epsilon_i \quad ; \quad \epsilon_i \sim \mathcal{N}(0, 1),
\]

- \( z_i \in \{1, \ldots, K\} \) hidden variable: the class label of the component generating \( x_i \)
- \( \beta_{z_i} \in \mathbb{R}^{p+1} \): regression coefficients of the sub-model \( z_i \)
- \( r_i = (1, f_i, \ldots, f_i^p)^T \): covariate vector in \( \mathbb{R}^{p+1} \)

\( z = (z_1, \ldots, z_n) \) is a hidden logistic process

\( z_i \sim \mathcal{M}(1, \pi_{i1}(w), \ldots, \pi_{iK}(w)); \) where

\[
\pi_{ik}(w) = p(z_i = k; w) = \frac{\exp(w_{k0} + w_{k1} f_i)}{\sum_{\ell=1}^K \exp(w_{\ell0} + w_{\ell1} f_i)},
\]

- \( w = (w_{10}, w_{11}, \ldots, w_{K0}, w_{K1}) \in \mathbb{R}^{2K} \) the parameter vector for the \( K \) logistic functions
Flexibility of the logistic transformation

Variation of $\pi_{ik}(w)$ in relation to $w$:

- Use the notation $w_k = (w_{k0}, w_{k1})^T = w_{k1}(\frac{w_{k0}}{w_{k1}}, 1)^T = \lambda_k(\alpha_k, 1)^T$

- Example of two components:

$$\alpha_1 = -2$$
Feature extraction

A regression model with a hidden logistic process

Flexibility of the logistic transformation

Variation of $\pi_{ik}(w)$ in relation to $w$:

- Use the notation $w_k = (w_{k0}, w_{k1})^T = w_{k1}(\frac{w_{k0}}{w_{k1}}, 1)^T = \lambda_k(\alpha_k, 1)^T$
- Example of two components:

\[ \alpha_1 = -2 \]

$\Rightarrow$ The parameter $\lambda_k$ controls the quality of transitions (smooth/abrupt) between the regimes
Flexibility of the logistic transformation

Variation of $\pi_{ik}(w)$ in relation to $w$:

- Use the notation $w_k = (w_{k0}, w_{k1})^T = w_{k1}(\frac{w_{k0}}{w_{k1}}, 1)^T = \lambda_k(\alpha_k, 1)^T$
- Example of two components:
  \begin{align*}
  \alpha_1 &= -2 \\
  \lambda_1 &= -5
  \end{align*}

⇒ The parameter $\lambda_k$ controls the quality of transitions (smooth/abrupt) between the regimes
⇒ The parameter $\alpha_k$ is directly linked to the frequency at the transition point
Parameter estimation by maximum likelihood

- Derived mixture density

\[
p(x_i; \theta) = \sum_{k=1}^{K} \pi_{ik}(w) \mathcal{N}(x_i; \beta_k^T r_i, \sigma_k^2)
\]

- Model parameters

\[
\theta = (w, \beta_1, \ldots, \beta_K, \sigma_1^2, \ldots, \sigma_K^2)
\]

- Log-likelihood of \( \theta \):

\[
L(\theta; x) = \sum_{i=1}^{n} \log \sum_{k=1}^{K} \pi_{ik}(w) \mathcal{N}(x_i; \beta_k^T r_i, \sigma_k^2).
\]

- Maximization of \( L(\theta; x) \) by a dedicated Expectation-Maximization (EM) algorithm [Dempster et al. 77].
Dedicated EM algorithm

Initialization: \[ \theta(0) \]

Expectation step: Compute the cond. expectation of the complete log-likelihood \[ Q(\theta, \theta(q)) = \mathbb{E}[L(\theta; x, z) | x, \theta(q)] = \frac{1}{n} \sum_{i=1}^{n} \sum_{k=1}^{K} \tau(q)_{ik} \log \pi_{ik}(w) + \sum_{i=1}^{n} \sum_{k=1}^{K} \tau(q)_{ik} \log N(x_i; \beta_k, \sigma_k^2) \]

Maximization step: Compute \[ \theta(q+1) = \text{arg max}_{\theta} Q(\theta, \theta(q)) \]

Maximization of \[ Q_2 \] w.r.t \{\beta_k, \sigma_k^2\} (\kappa = 1,...,K): Analytic solutions

Maximization of \[ Q_1 \] w.r.t \(w\): a multiclass weighted logistic regression problem \(\Rightarrow\) IRLS algorithm [Green 84, Jordan & Jacobs 94]
Dedicated EM algorithm

**Initialization:** $\theta^{(0)}$
Dedicated EM algorithm

Initialization: $\theta^{(0)}$

1. **Expectation step:** Compute the cond. expectation of the complete log-likelihood

$$Q(\theta, \theta^{(q)}) = E \left[ L(\theta; x, z)|x, \theta^{(q)} \right]$$

$$= n \sum_{i=1}^{n} K \sum_{k=1}^{K} \tau_{ik}^{(q)} \log \pi_{ik}(w) + n \sum_{i=1}^{n} K \sum_{k=1}^{K} \tau_{ik}^{(q)} \log \mathcal{N}(x_i; \beta_k^T r_i, \sigma_k^2),$$

where $\tau_{ik}^{(q)} = p(z_i = k|x_i; \theta^{(q)})$ is the posterior probability of the $k$th regime.
Dedicated EM algorithm

**Initialization: \( \theta^{(0)} \)**

1. **Expectation step:** Compute the cond. expectation of the complete log-likelihood

\[
Q(\theta, \theta^{(q)}) = E \left[ L(\theta; x, z) | x, \theta^{(q)} \right] 
\]

\[
= \sum_{i=1}^{n} \sum_{k=1}^{K} \tau_{ik}^{(q)} \log \pi_{ik}(w) + \sum_{i=1}^{n} \sum_{k=1}^{K} \tau_{ik}^{(q)} \log \mathcal{N}(x_i; \beta_k^T r_i, \sigma_k^2),
\]

where \( \tau_{ik}^{(q)} = p(z_i = k | x_i; \theta^{(q)}) \) is the posterior probability of the \( k \)th regime

2. **Maximization step:** Compute \( \theta^{(q+1)} = \arg \max_{\theta} Q(\theta, \theta^{(q)}) \)
Dedicated EM algorithm

**Initialization: \( \theta^{(0)} \)**

1. **Expectation step:** Compute the cond. expectation of the complete log-likelihood

\[
Q(\theta, \theta^{(q)}) = E \left[ L(\theta; x, z) | x, \theta^{(q)} \right]
\]

\[
= \sum_{i=1}^{n} \sum_{k=1}^{K} \tau_{ik}^{(q)} \log \pi_{ik}(w) + \sum_{i=1}^{n} \sum_{k=1}^{K} \tau_{ik}^{(q)} \log \mathcal{N}(x_i; \beta_k^T r_i, \sigma_k^2),
\]

where \( \tau_{ik}^{(q)} = p(z_i = k | x_i; \theta^{(q)}) \) is the posterior probability of the \( k \)th regime

2. **Maximization step:** Compute \( \theta^{(q+1)} = \arg \max_{\theta} Q(\theta, \theta^{(q)}) \)

   1. Maximization of \( Q_2 \) w.r.t \( \{\beta_k, \sigma_k^2\} \) (\( k = 1 \ldots, K \)): Analytic solutions
Feature extraction

EM algorithm

Dedicated EM algorithm

**Initialization:** $\theta^{(0)}$

1. **Expectation step:** Compute the cond. expectation of the complete log-likelihood

$$Q(\theta, \theta^{(q)}) = E\left[ L(\theta; x, z) | x, \theta^{(q)} \right]$$

$$= \sum_{i=1}^{n} \sum_{k=1}^{K} \tau_{ik}^{(q)} \log \pi_{ik}(w) + \sum_{i=1}^{n} \sum_{k=1}^{K} \tau_{ik}^{(q)} \log \mathcal{N}(x_i; \beta_k^T r_i, \sigma_k^2),$$

where $\tau_{ik}^{(q)} = p(z_i = k | x_i ; \theta^{(q)})$ is the posterior probability of the $k$th regime.

2. **Maximization step:** Compute $\theta^{(q+1)} = \arg \max_{\theta} Q(\theta, \theta^{(q)})$

   1. Maximization of $Q_2$ w.r.t $\{\beta_k, \sigma_k^2\}$ ($k = 1 \ldots, K$): Analytic solutions
   2. Maximization of $Q_1$ w.r.t $w$: a multiclass weighted logistic regression problem $\Rightarrow$ IRLS algorithm [Green 84, Jordan & Jacobs 94]
As in standard regression, given the estimated parameters, $x_i$ is approximated by its expectation:

$$\hat{x}_i = E(x_i; \hat{\theta}) = \int_{\mathbb{R}} x_i p(x_i; \hat{\theta}) dx_i$$

$$= \sum_{k=1}^{K} \pi_{ik}(\hat{w}) \hat{\beta}_k^T r_i$$

A sum of polynomials weighted by the logistic probabilities $\pi_{ik}(\hat{w})$'s

⇒ Adapted for a smooth or abrupt transitions between the regression models.

Segmentation

- The estimated class label $\hat{z}_i$ of $x_i$ can be computed by the rule:

$$\hat{z}_i = \arg \max_{1 \leq k \leq K} \pi_{ik}(\hat{w})$$
Case study:

- The impedance spectrums include 3 regimes which correspond to three behaviors of the stack → The number of regressive components is then set to $K = 3$
- The degree $p$ of the polynomial regression is set to 3 which is adapted to the different regimes in the curves
The obtained approximation
Linear regression model for the FC lifetime estimation

\[ LT_j = \alpha + b^T y_j + err_j \]

for the EIS \( j \) where:

- \( LT_j \): the duration time
- \( y_j = (a_j, \theta_j)^T \) features extracted from the real and the imaginary part
- \( (\alpha, b)^T \) the vector of regression coefficients
linear regression model for the FC lifetime estimation

\[ LT_j = \alpha + b^T y_j + err_j \] for the EIS \( j \) where:

- \( LT_j \): the duration time
- \( y_j = (a_j, \theta_j)^T \) features extracted from the real and the imaginary part
- \( (\alpha, b)^T \) the vector of regression coefficients

Mean error (in hours) of duration time estimation using different input descriptors:

<table>
<thead>
<tr>
<th></th>
<th>Training set</th>
<th>Test set</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real part (dim=1) ( a_2 )</td>
<td>181.40</td>
<td>194.02</td>
</tr>
<tr>
<td>Imag. part (dim=3) ( \beta_{23}, \beta_{32}, \beta_{34} )</td>
<td>137.06</td>
<td>153.53</td>
</tr>
<tr>
<td>Real + Imag. parts (dim=7) ( \beta_{21}, \beta_{23}, \beta_{24}, a_1, a_2, a_3, a_4 )</td>
<td>94.80</td>
<td>142.30</td>
</tr>
</tbody>
</table>
Duration time estimation obtained for the training set (left) and the test set (right):
Conclusion

- Supervised learning approach for Fuel Cell lifetime estimation from EIS measurements

- A probabilistic approach is used for feature extraction (from the imaginary part of the EIS)
  - The proposed model integrates a logistic process which makes possible to change smoothly within various possible regression models
  - Accurate modeling of the nonlinearities within the curves
  - Allows for automatically finding the three regimes corresponding to the behaviours of the stack
Thank you!