Factorization Machines

Ray Bosman – Senior Developer

DERIVCO

MARCH 2020

FM Overview

- Simon Funk, Funk MF
- Imputing missing values

Model

There are 2 matrices that have a dot product yielding a matrix R where the difference between R and our data is at a minimum

$$\begin{bmatrix} 1.2 \\ 1.3 \\ 1.5 \\ 2.2 \\ 1.6 \\ 1.0 \\ 1.4 \\ 1.4 \\ 1.6 \end{bmatrix} * \begin{bmatrix} 2 & 1.7 & 2.2 & 1.3 & 1.7 \end{bmatrix} \approx \begin{bmatrix} 2.3 & 2.0 & 2.6 & 1.5 & 2.0 \\ 2.6 & 2.3 & 3.0 & 1.7 & 2.2 \\ 2.9 & 2.6 & 3.3 & 1.9 & 2.5 \\ 4.4 & 3.8 & 5.0 & 2.9 & 3.7 \\ 3.1 & 2.7 & 3.5 & 2.0 & 2.6 \\ 2.0 & 1.7 & 2.2 & 1.3 & 1.7 \\ 2.7 & 2.4 & 3.1 & 1.8 & 2.3 \\ 2.7 & 2.4 & 3.1 & 1.8 & 2.3 \\ 3.2 & 2.8 & 3.7 & 2.1 & 2.7 \end{bmatrix}$$

Why would a derived matrix be useful?

Model

If a matrix exists that looks similar to our data, surely we would want to find It and borrow from it?

$$R = \begin{bmatrix} 2.3 & 2.0 & 2.6 & 1.5 & 2.0 \\ 2.6 & 2.3 & 3.0 & 1.7 & 2.2 \\ 2.9 & 2.6 & 3.3 & 1.9 & 2.5 \\ 4.4 & 3.8 & 5.0 & 2.9 & 3.7 \\ 3.1 & 2.7 & 3.5 & 2.0 & 2.6 \\ 2.0 & 1.7 & 2.2 & 1.3 & 1.7 \\ 2.7 & 2.4 & 3.1 & 1.8 & 2.3 \\ 2.7 & 2.4 & 3.1 & 1.8 & 2.3 \\ 3.2 & 2.8 & 3.7 & 2.1 & 2.7 \end{bmatrix}$$

$$\text{Our data} = \begin{bmatrix} 2 & ? & ? & ? & 2 \\ ? & ? & 3 & 2 & ? \\ 3 & ? & ? & ? & ? \\ 4 & 4 & 5 & 3 & 4 \\ ? & ? & ? & 2 & ? \\ 2 & 2 & 2 & 1 & 1 \\ 3 & ? & 3 & ? & ? \\ ? & ? & 3 & ? & 2 \\ 3 & 3 & ? & 2 & ? \end{bmatrix}$$

Rating Imputation

Before imputing

Service

Users

2	?	?	?	2
?	?	3	2	?
3	?	?	?	?
4	4	5	3	4
?	?	?	2	?
2	2	2	1	1
3	?	3	?	?
?	?	3	?	2
3	3	?	2	?

After imputing

Service

Users

2	**2**	**3**	**2**	2
3	**2**	3	2	**2**
3	**3**	**3**	**2**	**3**
4	4	5	3	4
3	**3**	**4**	2	**3**
2	2	2	1	2
3	**2**	3	**2**	**2**
3	**2**	3	**2**	2
3	3	**4**	2	**3**

How do we find these factors?

How do we find these factors?

$$\begin{bmatrix} ? \\ ? \\ ? \\ ? \\ ? \\ ? \\ ? \\ ? \\ ? \end{bmatrix} * \begin{bmatrix} ? & ? & ? & ? & ? \end{bmatrix} :$$

Gradient Descent

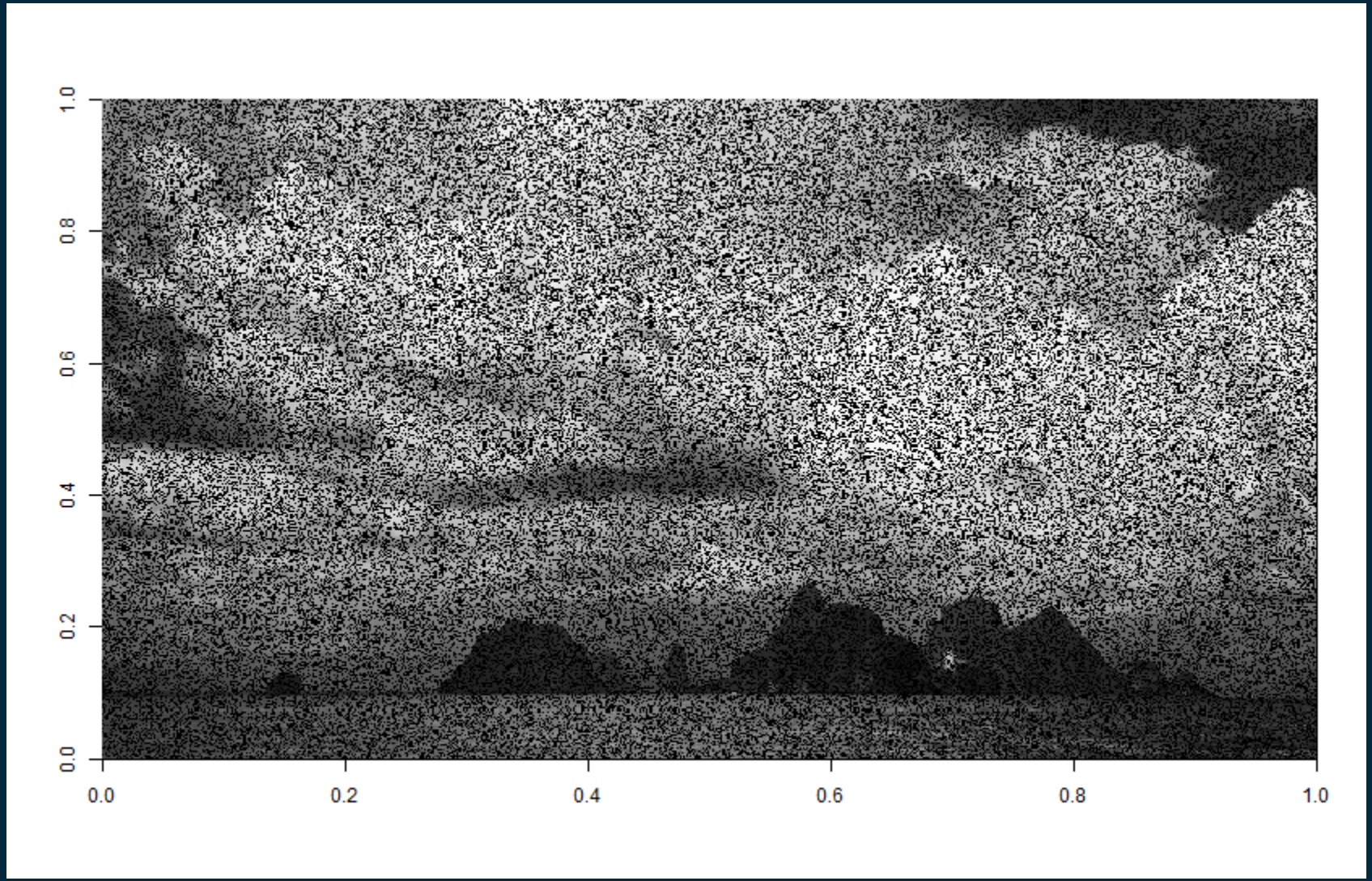
$$x_{ik} = x_{ik} - \alpha(-2(R_{ij} - x_i y_j)y_{kj})$$

$$y_{kj} = y_{kj} - \alpha(-2(R_{ij} - x_i y_j)x_{ik})$$

Gradient Descent in R

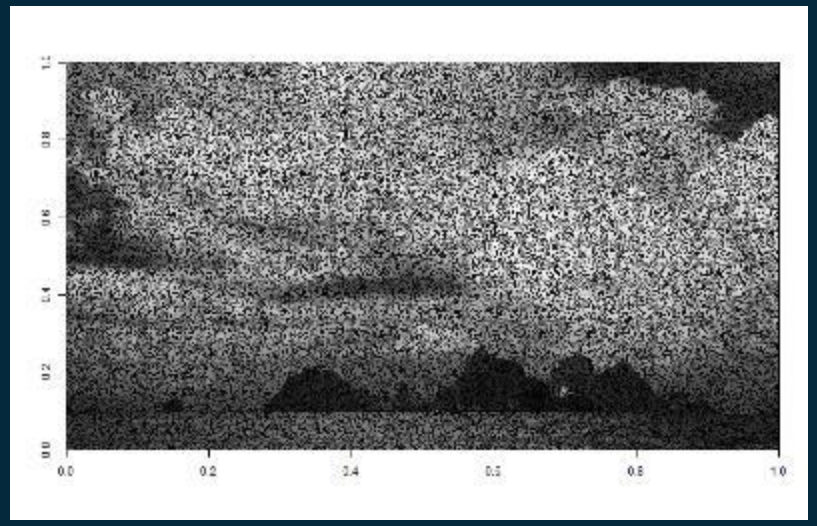
```
X[i,k] <- X[i,k] - alpha * (-2 * (eij) * Y[k,j])
Y[k,j] <- Y[k,j] - alpha * (-2 * (eij) * X[i,k])
```

Pixel imputation : Sample 1 - corrupted



Pixel imputation : Sample 1 - generated

Pixel imputation : Sample 1 - corrected

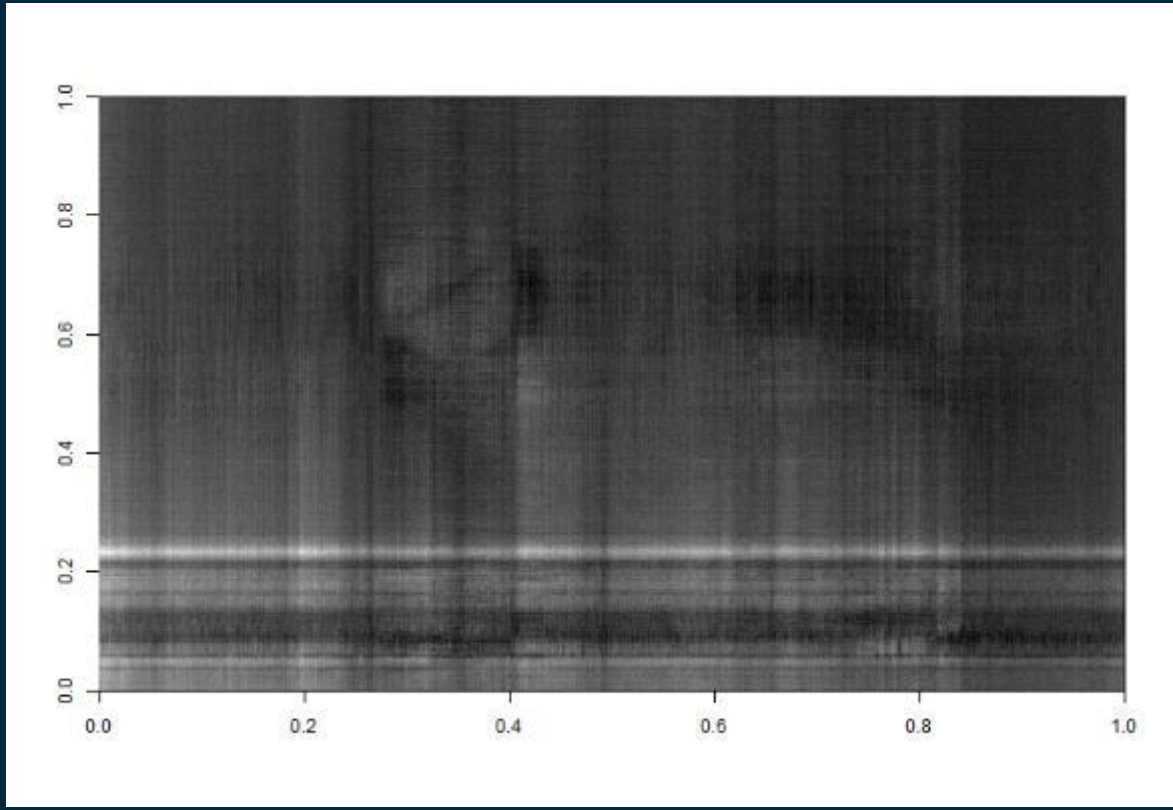


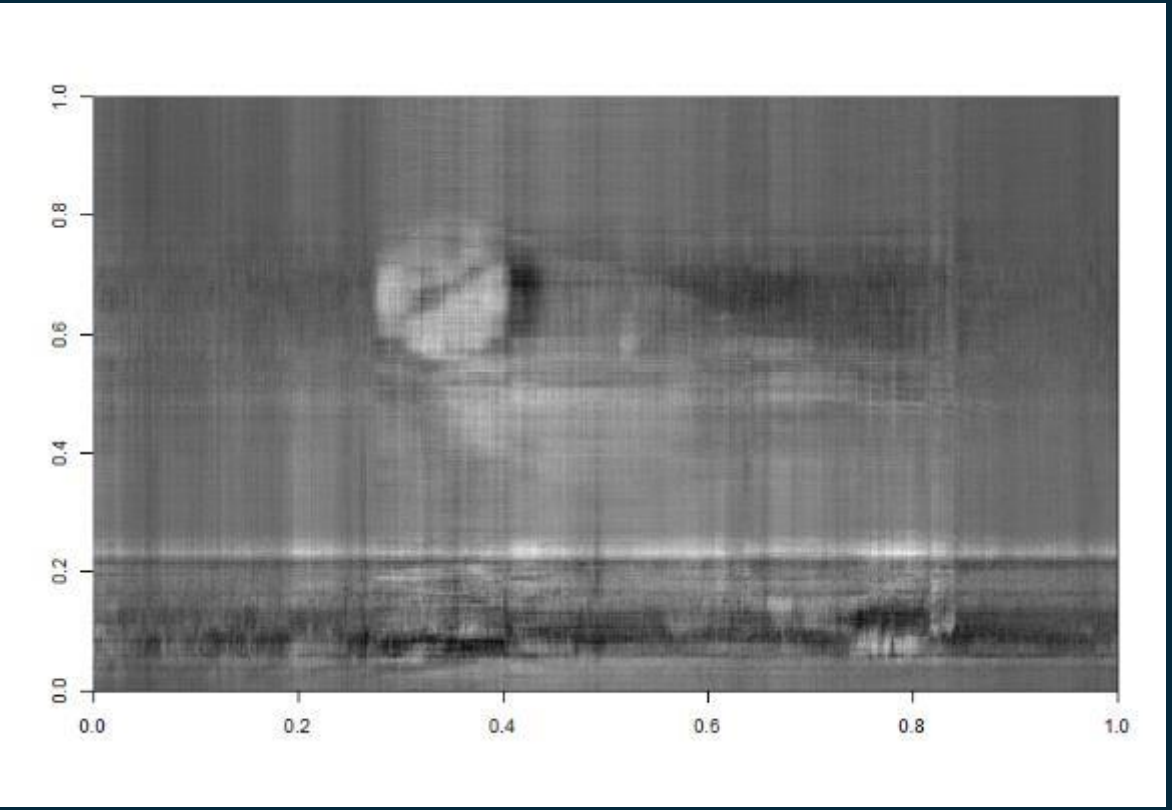
+

=

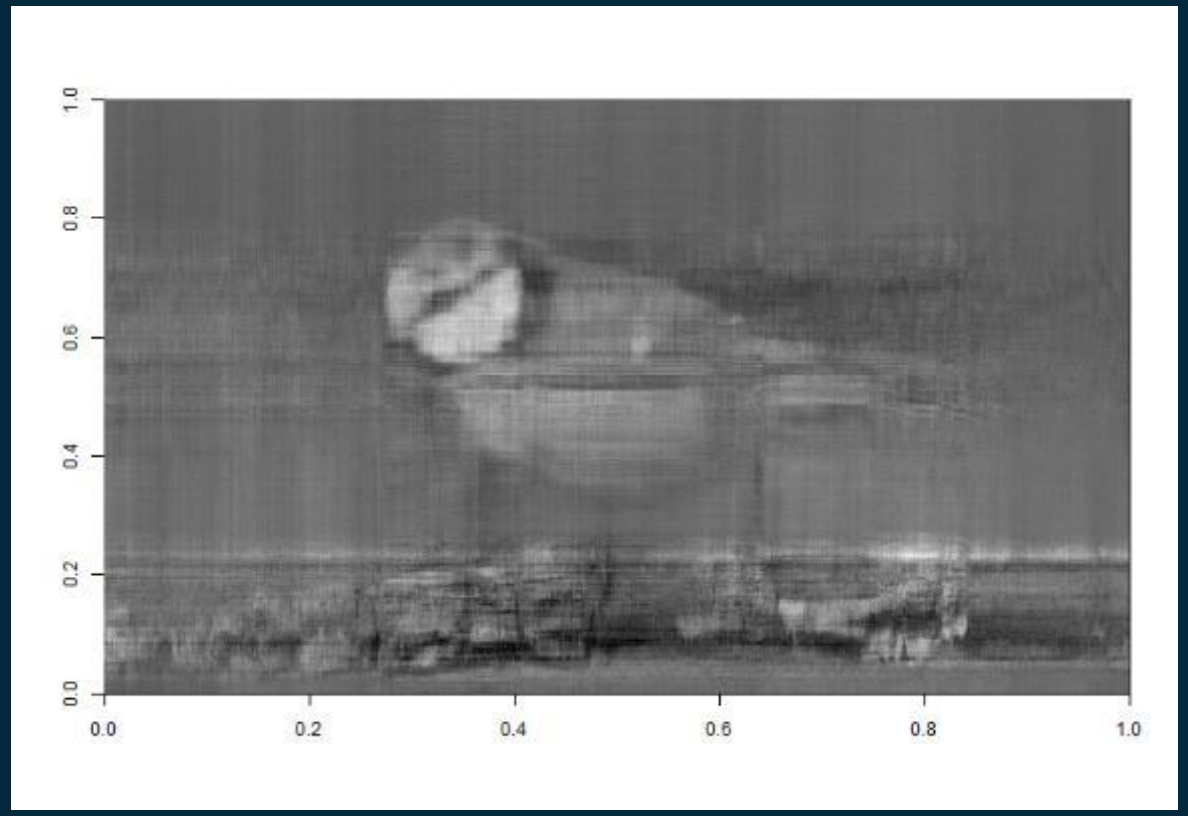
Visual of incremental progression

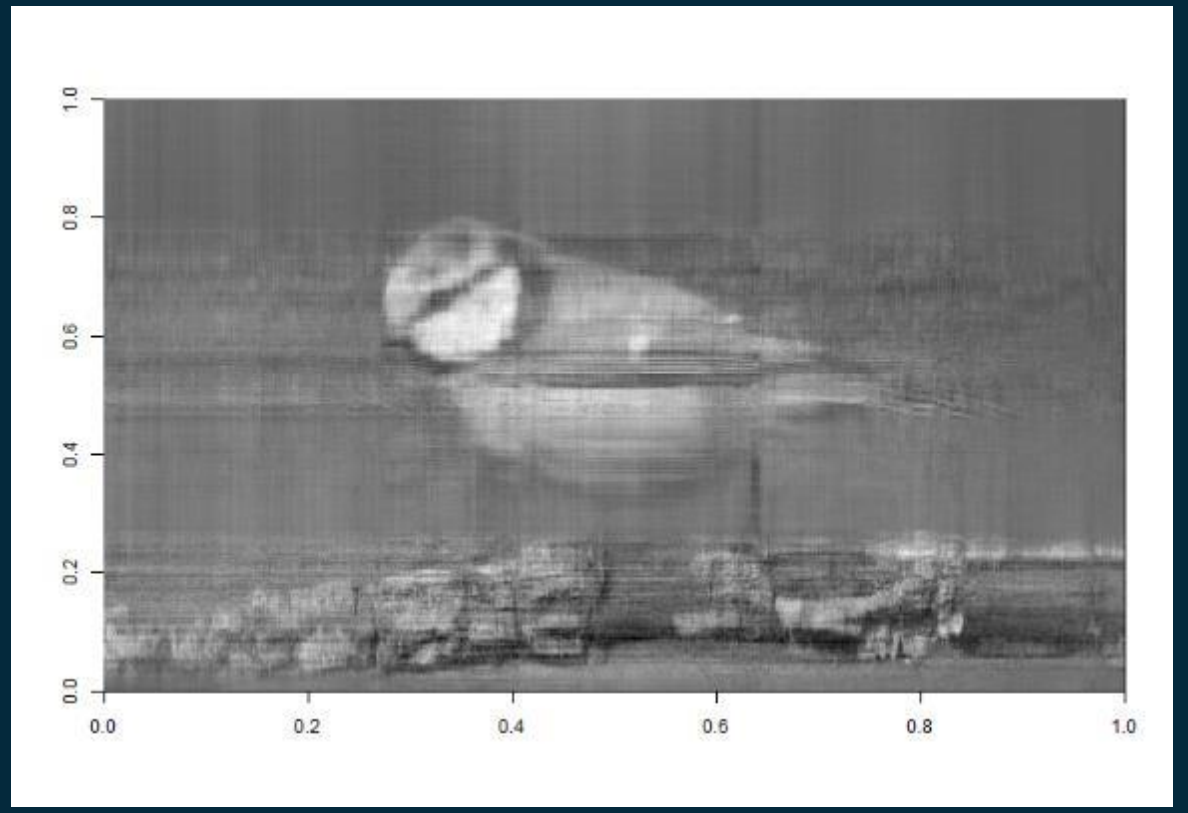
Incremental progression





Incremental progression





Incremental progression

Incremental progression

Kudos to Albert Au Yeung

- Albert Au Yeung
  Machine Learning Engineer, Hong Kong

- Python Funk MF tutorial code

- Extended for images and parallelism

Contact

- **Email:** currieroad@gmail.com
- **LinkedIn:** www.linkedin.com/in/ray-bosman
- **Github:** https://github.com/SledgeHardy/