Tampere University

ICPR 2020

AARHUS UNIVERSITY

25th International Conference on Pattern Recognition

Data Normalization for Bilinear Structures in High-Frequency Financial Time-series

Dat Thanh Tran, Moncef Gabbouj and Alexandros Iosifidis

Main Contributions

- We proposed a learnable data-normalization method for multivariate time-series
- This method takes into account the tensor nature of multivariate time-series
- Outperform other normalization schemes in stock movement prediction

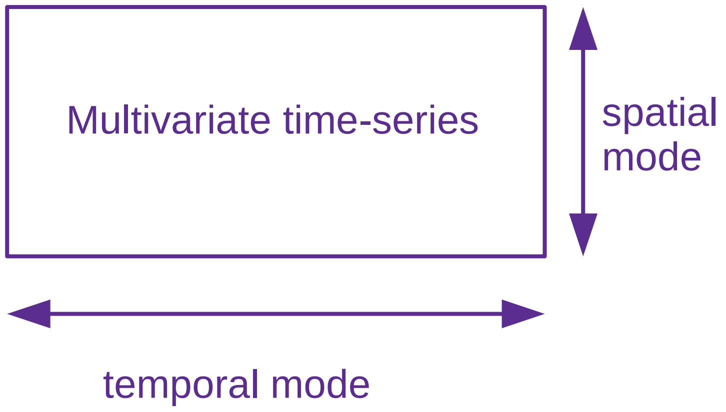
Tampere University

ICPR 2020

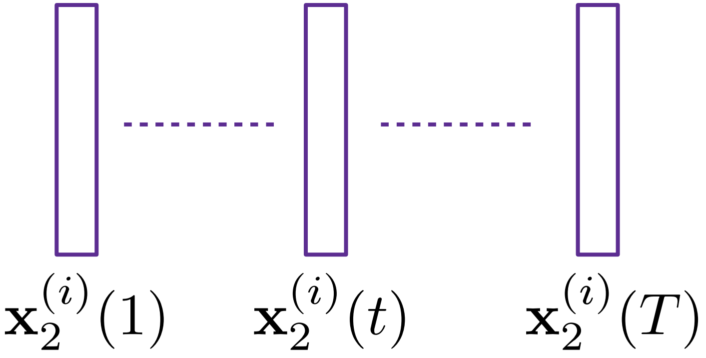
AARHUS UNIVERSITY

Bilinear Normalization (BiN)

- Designed for multivariate time-series, which are tensors (2D matrices)
- Normalize multivariate time-series along both the temporal and spatial/channel/feature modes.



Temporal Mode Normalization



$\mathbf{x}_2^{(i)}(t)$

- superscript (i) denotes the i-th series in training set
- (t) denotes the index of the slice
- subscript 2 denotes the slices in temporal mode

$$\bar{\mathbf{x}}_2^{(i)} = \text{mean}\big(\mathbf{x}_2^{(i)}(t)\big), \forall t = 1, \dots, T$$

mean of the slices in temporal mode

$$\sigma_2^{(i)} = \text{std}\big(\mathbf{x}_2^{(i)}(t)\big), \forall t = 1, \dots, T$$

std of the slices in temporal mode

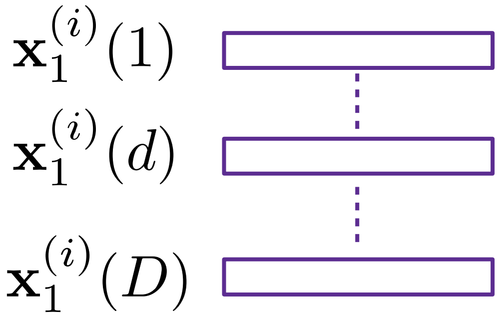
$$\mathbf{Z}_2^{(i)} = \left(\mathbf{X}^{(i)} - \bar{\mathbf{x}}_2^{(i)}\mathbf{1}_T^{\mathrm{T}}\right) \oslash \left(\sigma_2^{(i)}\mathbf{1}_T^{\mathrm{T}}\right)$$

Z-score normalization step

$$\tilde{\mathbf{X}}_2^{(i)} = \left(\gamma_2\mathbf{1}_T^{\mathrm{T}}\right) \odot \mathbf{Z}_2^{(i)} + \beta_2\mathbf{1}_T^{\mathrm{T}}$$

scale and shift with γ_2 and β_2

Spatial Mode Normalization



$\mathbf{x}_1^{(i)}(d)$

superscript (i) denotes the i-th series in training set

(d) denotes the index of the slice

subscript 1 denotes the slices in spatial/channel mode

$$\bar{\mathbf{x}}_1^{(i)} = \mathrm{mean}\big(\mathbf{x}_1^{(i)}(d)\big), \forall d = 1, \ldots, D$$

mean of the slices in temporal mode

$$\sigma_1^{(i)} = \mathrm{std}\big(\mathbf{x}_1^{(i)}(d)\big), \forall d = 1, \ldots, D$$

std of the slices in temporal mode

$$\mathbf{Z}_1^{(i)} = \big(\mathbf{X}^{(i)} - \mathbf{1}_D(\bar{\mathbf{x}}_1^{(i)})^{\mathrm{T}}\big) \oslash \big(\mathbf{1}_D(\sigma_2^{(i)})^{\mathrm{T}}\big)$$

Z-score normalization step

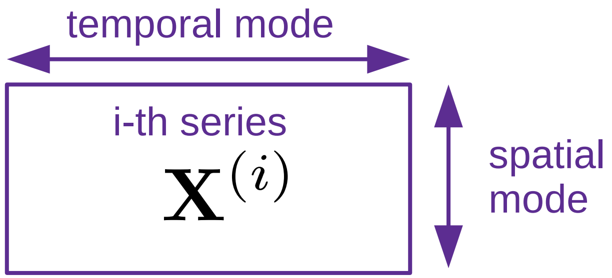
$$\tilde{\mathbf{X}}_1^{(i)} = \big(\mathbf{1}_D\gamma_1^{\mathrm{T}}\big) \odot \mathbf{Z}_1^{(i)} + \mathbf{1}_D\beta_1^{\mathrm{T}}$$

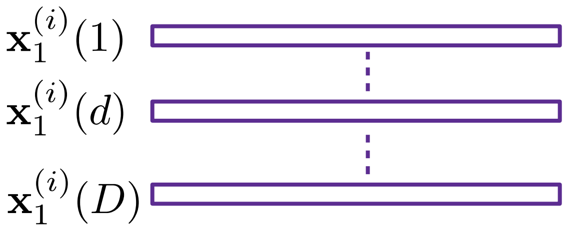
scale and shift with γ_1 and β_1

Tampere University

ICPR 2020

AARHUS UNIVERSITY





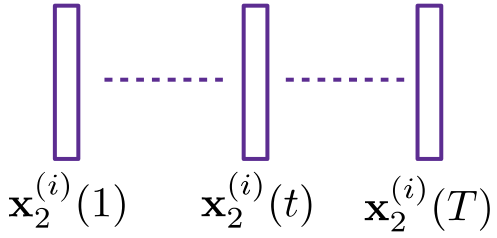
$$\bar{\mathbf{x}}_1^{(i)} = \mathrm{mean}\big(\mathbf{x}_1^{(i)}(d)\big), \forall d = 1, \dots, D$$

$$\sigma_1^{(i)} = \mathrm{std}\big(\mathbf{x}_1^{(i)}(d)\big), \forall d = 1, \dots, D$$

$$\mathbf{Z}_1^{(i)} = \big(\mathbf{X}^{(i)} - \mathbf{1}_D(\bar{\mathbf{x}}_1^{(i)})^{\mathrm{T}}\big)\big(\mathbf{1}_D(\sigma_2^{(i)})^{\mathrm{T}}\big)$$

$$\tilde{\mathbf{X}}_1^{(i)} = \big(\mathbf{1}_D\gamma_1^{\mathrm{T}}\big) \odot \mathbf{Z}_1^{(i)} + \mathbf{1}_D\beta_1^{\mathrm{T}}$$

spatial mode normalization



$$\bar{\mathbf{x}}_2^{(i)} = \mathrm{mean}\big(\mathbf{x}_2^{(i)}(t)\big), \forall t = 1, \dots, T$$

$$\sigma_2^{(i)} = \mathrm{std}\big(\mathbf{x}_2^{(i)}(t)\big), \forall t = 1, \dots, T$$

$$\mathbf{Z}_2^{(i)} = \big(\mathbf{X}^{(i)} - \bar{\mathbf{x}}_2^{(i)}\mathbf{1}_T^{\mathrm{T}}\big)\big(\sigma_2^{(i)}\mathbf{1}_T^{\mathrm{T}}\big)$$

$$\tilde{\mathbf{X}}_2^{(i)} = \big(\gamma_2\mathbf{1}_T^{\mathrm{T}}\big) \odot \mathbf{Z}_2^{(i)} + \beta_2\mathbf{1}_T^{\mathrm{T}}$$

temporal mode normalization

$$\tilde{\mathbf{X}}^{(i)} = \lambda_1\tilde{\mathbf{X}}_1^{(i)} + \lambda_2\tilde{\mathbf{X}}_2^{(i)}$$

Tampere University

ICPR 2020

AARHUS UNIVERSITY

- Problem: stock movement prediction
- Horizon corresponds to different time in the future (H=10, 20, 50)
- F1 as the main performance measure (imbalanced dataset)

Models	**Accuracy %**	**Precision %**	**Recall %**	**F1 %**
Prediction Horizon $H = 10$				
CNN[25]	-	50.98	65.54	55.21
LSTM[26]	-	60.77	75.92	66.33
C(BL) [1]	82.52	73.89	76.22	75.01
DeepLOB [2]	84.47	84.00	84.47	83.40
DAIN-MLP [20]	-	65.67	71.58	68.26
DAIN-RNN [20]	-	61.80	70.92	65.13
C(TABL) [1]	84.70	76.95	78.44	77.63
BN-C(TABL)	79.20	68.48	72.36	66.87
BiN-C(TABL)	86.87	80.29	81.84	**81.04**
Prediction Horizon $H = 20$				
CNN[25]	-	54.79	67.38	59.17
LSTM[26]	-	59.60	70.52	62.37
C(BL) [1]	72.05	65.04	65.23	64.89
DeepLOB [2]	74.85	74.06	74.85	72.82
DAIN-MLP [20]	-	62.10	70.48	65.31
DAIN-RNN [20]	-	59.16	68.51	62.03
C(TABL) [1]	73.74	67.18	66.94	66.93
BN-C(TABL)	70.70	63.10	63.78	63.43
BiN-C(TABL)	77.28	72.12	70.44	**71.22**
Prediction Horizon $H = 50$				
CNN[25]	-	55.58	67.12	59.44
LSTM[26]	-	60.03	68.58	61.43
C(BL) [1]	78.96	77.85	77.04	77.40
DeepLOB [2]	80.51	80.38	80.51	80.35
DAIN-MLP [20]	-	-	-	-
DAIN-RNN [20]	-	-	-	-
C(TABL) [1]	79.87	79.05	77.04	78.44
BN-C(TABL)	77.16	75.70	75.04	75.34
BiN-C(TABL)	88.54	89.50	86.99	**88.06**

Thank You for listening!