

Á

Á

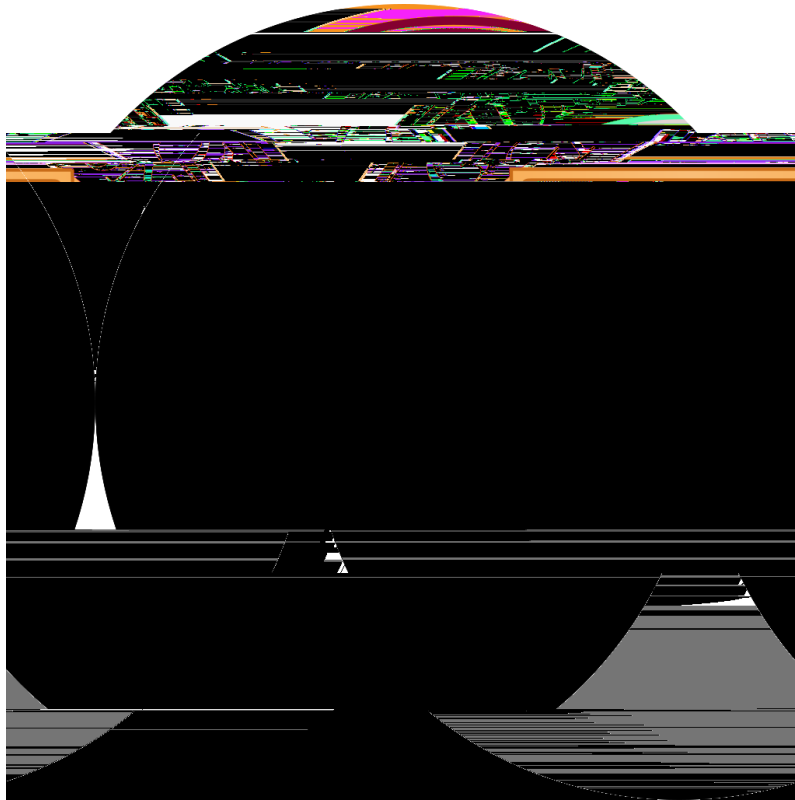
Á

Á

ŒŒ æ&@Œ ^Š^æ} ā * ŒŒ]|[æ&@Œ ŒŒ Œ { æ^âŒ!æâŒ * Á

Boston College Computer Science Senior Thesis

Á



Á

Á

Á

Œā * Š~ Á

ŒŒŒ [ŒŒŒ * Œ ŒŒŒŒ : Á

T æ ŒŒŒŒŒ Œ Á

Á

Á

Á

Á

Á

Á

OE • d a&cÁ

Á

V@Á a &ãÁ a\^cÁ a& {]|^cÁ a/á^} a &ãÁ

FEŮ d[ā ~ & cā } Á

Á

CE Á - Ą GEFGĹA^][!d[{ Á [

GEÁE d { æ^âÁ/æâ * Á

Á

GEÁUç^!çã, Áæ âÁ^!{ • Á

Á

V@ÁUç & Á æ\^óÁÁ æ\^q|æ^Á @!^Á @e^Á-Á-Á~à|æÁ {] æ â•Áæ^Áæ^âÉÁ
& {] æ ^Á^& { ^•Á~à|æÁ @} Áæ æ^•Áæ ÁQæáU~à|æÁ-^!ç * É!Á { { [] ^Á } [, } Áæ ÁDUÉÁ
, @æÁ ^æ•Á@æç ç^•ç!•Á [!á, æ^Áæ^Áæ|Á Á~^Áæ âÁæ^Á @e^Á-Á-Áç & Á@Á {] æ ^ÉÁ
V@•^Á @e^Á^!^•^} óæó, }^!• @ Á Á@Á {] æ Éæ âÁ@áÁ|æ^Á^!^•^} ó @æç ç^•ç!•Á
à|áç^ÁÁ

[!â^!•Áç^!^âæÁ Á^^\Á dââæ Á | [-ãËZ | Áe d { æ^âÁæãã * Á^•c\ { •Ë^&ě •^Á -Á@ãÁ
•^•c\ { æãÁ æ | ^Ë@Áæãã * Á^~^ } & Á^^â•Á Á^Á] ^&ãããâ^Á@ãÁ^ç^ [] ^! •Á^ { | ^Áç^ } Á
â^ã } ã * Á@Á@ { Ë^!æãã * Á^~^ } & Á@•Áç^!^ Á æ^Áã * ^Á^! { Á } & ÁãÁ^ã ^Ë @&Á ^ã • Á
à^ Áã áÁ@ | áÁ | ^ç^! Ë Á } & Áç^!^ Á æ [•^&] áÁ | Áã @Ë^~^ } & Á!æãã * Áç [!ã@ •ËÁÁ
!^~ | Ë [

•••{ Á ã ã ã ^•Á { [ç] •Á @ [~ * @ ~ Á @ Á ç ç ç * Á ! [& •• Ë] • ~ ! ^•Á @ ç ç ã & ç | ç ^ Á Á ç ç ç ^ á Á
ç @ [~ * @ ç [| ç ç ^ Á ç \ ^ ç] } á ç ç } • Ë ç } á Á ç [, Á • Á ç Á ç @ ç ^ Á] • ã ç } & Ë Á

V @ Á @ | Á ç ç ç ç ^ Á - Á ç * [| ç @ ç Á ç ç ç * Á ç ^ | Á @ { ç Á ç ç ^ • Á Á @ Á ç ç ç Á ç Á ç ç ç • Á @ Á

•ç ;

Financial Time Series Forecasting with Machine Learning Techniques: A Survey

Á

QÁ@Á{ ä Á Á • ä * Á ä@ ^ Á ä } ä * Á & @ ä ~ ^ • Á Á | ^ ä ä d & Á ä ^ d [ç ^ { ^ } • Ä
Financial Time Series Forecasting with Machine Learning Techniques: A Survey
& @ • ä ^ Á ç ^ ç ä , Á Á ^ & ^ } ä ^ ç ^ [] { ^ } • Á ~ d Á Á Á ~ ä ä ä } • Á @ Á ~ | ç ^ Á & ç ^ • Ä Ç Á Á @ { Á
~ • ^ Á Ä ä

HÉT æ&@ ^Š^æ} ā * Á ^c@ à•Á

Á

HÉÁ!ā~Á d[à~ &cā } Á

Á

Qeā Á~!ā * Ē} ^Á -Á@Á!^æ•o&[{]~c!Á&a} cā•Ē Á@{ æ Ác d!^Á āÁæ@!Á -ÁcāāāÁ
ā c||ā^} &Ē Á@Á æ^!Á@ [{]~cā * Á æ&@ ^!^Á āÁc||ā^} &~Ē+Ēe\•Á@Áæ [~•Á~^•cā } KÁ
%ā Á æ&@ ^•Á@ \Ñ+ÁM} cĒ āæĒ ^ÁeÁ@ { æ Ác!Á } QÁ [, Á@Á• , ^!Á Á@Á~^•cā } Ē~cā
çæāā

β_j is the coefficient for the j -th predictor in the linear model. The goal is to find the best values for β_j that minimize the sum of squared residuals.

The least squares method finds the values of β_j that minimize the sum of squared residuals.

$$\sum_{i=1}^N \left(y_i - \sum_j \beta_j x_{ij} \right)^2 + \lambda \sum_j \beta_j^2$$

Á

ÁÁÁÁÁÁ

The least squares method finds the values of β_j that minimize the sum of squared residuals. This is done by taking the derivative of the sum of squared residuals with respect to each β_j and setting it equal to zero. This results in a system of linear equations that can be solved for the values of β_j .

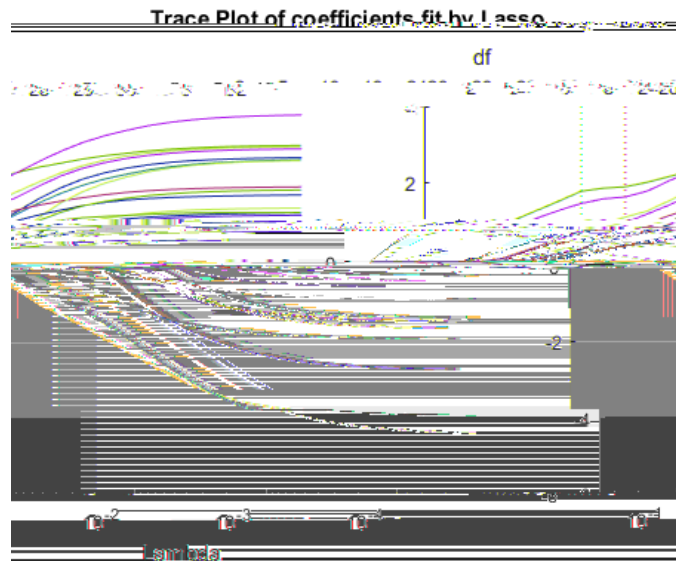
The least squares method finds the values of β_j that minimize the sum of squared residuals.

$$\sum_{i=1}^N \left(y_i - \sum_j \beta_j x_{ij} \right)^2 + \lambda \sum_j |\beta_j|$$

Á

ÁÁÁÁÁÁ

The least squares method finds the values of β_j that minimize the sum of squared residuals. This is done by taking the derivative of the sum of squared residuals with respect to each β_j and setting it equal to zero. This results in a system of linear equations that can be solved for the values of β_j .



Á

The least squares method finds the values of β_j that minimize the sum of squared residuals. This is done by taking the derivative of the sum of squared residuals with respect to each β_j and setting it equal to zero. This results in a system of linear equations that can be solved for the values of β_j .

P[, ^ç!Éáá^] Á^Á ä@ ä ^ Áæ^ Á d|Á) á•Á Á ç!-ä@ ä@!^|!^Á Á•Á& |æ: Á
ä Á -} ÁæÁ•Á@



Á

øã~!^Á KÖ Á cæ]|^Á -ÜUÔ&~!ç^Áæ)ã!~ Á] ã æ!@^•@|ãÁ

Á

ÜUÔ&~!ç^Á} Á^}^!æ! Á•^ã! &{]æ^Á^!+!{ æ&•Á&[••Áã^!^} ó&æ•ã!•Ë^Á
&{]æã* Á&@æ•ã!ç^ÁVÔ(ç^Áæ)ã!&~!ç^ÁD! -ÜUÔ&~!ç^ÁÉÓ~!ç^Á@ Áæ^É Á•^Á@Á
ÜUÔ&~!ç^Á Áã

| ÈÚ!^á&ç } Á [á^|KQ áã ä~ æÁ] | [æ&@Á

Á

| ÈÁ^c@ á [| * ^ Á

Á

V@Áá•á ^c@ áÁ Á^ç^ [| ÁÁ d & Á |æ^Á!^á&ç } Á [á^|Á Á @Á ^Á&ç|áÁ Á áã ä~ æÁ
æ] | [æ&@Á @Á ^á • Á@Á@Á [á^|Á • Á@ d |æ^Á!^ç | { æ & Á - ÁÁ d & Á ^|Á Á |^á&ç Á
~ ç |^Á |æ^Á [ç { ^ } • ÈV@Á^á Á@Á [| á * ÁÁQ, Á@Á d & Á@ d |æ^Á [ç • Á Á Èæ Á
, á ä

I EGÖææÁ

Á

V@ÁææÁ ^Á•^ÁÁ Á ^àÁ&æ ^áÁ{ ÁæQ[Áæ æ &É } æ•Áæb•c'á&[•q*Á
]!æ^Á!Á@Áç &•Á ÆBÚÁ ∈Á[{ Áæ ÁGEFI Á Á'áÁÉGEFÍ ÉY ^ÁæÁ[ó[Á'!c@!ÁæÁ Á
à^ææ•^Á!æ•Á [!^Ácæ ÁÁ^æ•Á|áÁÁ^|Á Á^Á|!^çæ c'á&[•q*Á!æ^Á@ÁæÁ Á
&[•q*Á!æÉ @ÁÁ@Áæc'áæq*Á!æ^Á!^Á æ\^c&[•^•ÁæÁ Á { Éæb

| ÈÁÛ^•~ |◦ Á

Á

| ÈÈÁÒçæ~ æā } Á ^d æÁ

Á

V@!^Áæ^Áç [Á æ•Á@æ, ^Á•^Áç Á@Á@•ã Áç çæ æ^Á@ Á^!-{| æ &^

| | | | | | | |
|----------|----------------------|------------------|------------------|---------------------------|-----------------|-----------------|
| Á | Š[* ā āŪ^*!^••ā } Á | | | Üā^*^Š[* ā āŪ^*!^••ā } Á | | |
| V@^•@ áÁ | VÚÁ | VÞÁ | VÜÁ | VÚÁ | VÞÁ | VÜÁ |
| €Ě Á | €Ě €€ġĚ Ĩ ĠÁ | €Ě FFI ġĚ Ě FDÁ | €Ě Ě Ĩ ġĚ Ě GFDÁ | €Ě €HġĚ Ĩ EDÁ | €Ě Ě Ĩ ġĚ Ĩ EDÁ | €Ě Ě Ě ġĚ Ĩ ĠÁ |
| €Ě Ĩ Á | €Ě FJġĚ Ĩ Ĩ DÁ | €Ě Ě Ĩ ġĚ Ĩ Ĩ DÁ | €Ě F Ĩ ġĚ Ĩ Ĩ DÁ | €Ě Ĩ Ĩ ġĚ Ĩ Ĩ DÁ | €Ě J Ĩ ġĚ Ĩ EDÁ | €Ě Ġ Ĩ ġĚ Ĩ FDÁ |
| | | | | | | |

Í ËÚ!^ã&ç } Á [ã^|Á^&ç :Á] : [æ&@Á

Á

Í ÈÁ^ç@ ã [| * ^ Á

Á

V@Á^á-á@Á^&ç :Á] : [æ&@Á Á , Á~!Á [ã^|Á Á [\ÁÁá *^!Á æ

Ø[{ Á@·^ÁãÁ [â^·Ëÿ~!ÁææÁ)áÁ [Á)·^ { à^Ë ^Á ãÁ æ Á@Á] Á^!-!{ ã*Á }^·Á
ç Áæá c·cÁ } Á^æÁ~ ðË Ëæ]|^ÁææÁ Á@Á^cÁ^&cÁ } ËÁ

Á

Á

í ËÖæ

Í ÈÄÛ^•~ |◦Á

Á

Í ÈÈÁi^ ^Ûæ•Á

Á

Ùã ãæÁí Á@Áã] | [æ&@Á @Á | ^çã ~ •Á ^&çã } Èÿ ^Á] | ãÁ @Á ææ ^óã ç Á €Á ^!& } óí æã ã * Á
æ ã Á €Á ^!& } óí •ç * ÈÛã &Á æã ã * Á !!! | Á ^æ •Á [çã * Á @ } Á] | { ^ } ç * Á @ Á

| | | | |
|---|---|---|--|
| Á | $V_i \sim \hat{A} \cup \cdot \tilde{a} \tilde{a} \hat{A}$ | $V_i \sim \hat{A} \wedge \hat{A}^* \tilde{a} \tilde{a} \hat{A}$ | $V_i \sim \hat{A} \tilde{a} \tilde{a}$ |
| | | | |
| | | | |
| | | | |

æ Á | | , • Á Á @ Á | ä ä æ Á æ Á Ö æ Á • Á | ^ ä æ Á] Æ @ Á ä æ Á [ä | Á æ Á • Á @ Á | ^ ä æ Á } Æ ~ ö Á @ Á
 [| ä ä æ Á [ä | Á | ^ ä æ Á • Á | , } Æ @ Á ä æ Á [ä | Á } | Á æ Á • Á @ Á æ Á | ^ ä æ Á } Á Á @ Á | : ä ä æ Á Á - Á
] | ^ ä æ Á * Á Á ^ * æ Á ^ Á ^ c | } Á ä ^ } Á Á @ Á | ä ä æ Á [ä | Á ç & ^ ä • Á @ Á | ç æ Á @ Á • @ | ä Á | { Á @ Á
 Ü Ü Ö Á & | ç ^ Æ Á

Á

| Á | V ^ Á • ä ä ^ Á | V ^ Á • ^ * ä ä ^ Á | V ^ Á Ü æ Á |
|-----------------------------------|-------------------------|-------------------------|-------------------------|
| W ä ç Á | € È J I J Ä È È G H Á | € È € H Ä È È G H Á | € È Í F H Ä È È Í Í D Á |
| Ö ^ * ^ Á | € È Í J Í Ä È È Í Í D Á | € È Í Í Í Ä È È Í Í D Á | € È G F € Ä È È H Í D Á |
| Q + { ä ä } Á ^ & @ [* ^ Á | € È F F Í Ä È È G H Á | € È € Í Í Ä È È J Í D Á | € È € J Í Ä È È H Í D Á |

Á ä | ^ Á | ç | ä ^ Á | ^ Á æ Á • Á - Á æ Á Ö æ Á • Á ä @ Ü Ü Ö Á ä æ Á • ä Á } Á ä - Á | ^ } ö ^ & ç | • Á

Á

V @ Á | ^ Á ^ * ä ä ^ Á æ Á • Á | | : ç ^ Á | ä @ ^ Á • ä * Á @ Á | [È ç] Á ^ | Æ | ç ç * @ Ä ^ Á | ä ç ^ Á @ Á
 ä] | | ç ^ { ^ } ö æ Á [ö ^ Á ä] ä ä æ Á Æ Á Á [[ä | Á Á ^ Á [Á ä } • Á - Á ç | - ä ä * Æ Á
 Þ ^ ç Æ Á ^ Á [| \ Á Á @ Á | ^ Á æ Á • Á | { Á } } ä * Á @ Á | [Á] • ^ { ä | Á ^ ç @ ä • Æ È æ ä Æ Ü



| | | | |
|--|--|--|--|
| | | | |
| | | | |

P[, ^ç^!Ë[{ ^Á Á@Á^ Áæ^ Á ã@Ë çç^ Áæ^ Á@Á ÁËFÁæ } [cÁ^b&Á@Á^ ||Á@] [c@•ã ÁæÁ
FÁ Áã ããææ &^Áç^|ËË [Ë@ÁË^•cÁç] -ã{ Á^!Á!^çç^ •Áã äã * Á@ÁÛXT Áã äÁ æç^ Áóæ^ Á
@ç^ Á^c^!Á] •ã^Á!^ããç } •Á@ËÁ [, } •ã^Ëã &^Á@Á^ Á^ Á^* æç^ Áæ^ Á^ [{ Á [cÁ çç Áã äÁ
ã { { æç } Á &@ [[^ Á^ Áç! •ÁçÁ Á^ Áã^!^ } c@ç Áã ä [{ Á^ Á^•ã * ËÁ

0ËË ä

| | | | |
|----------------------------------|-------------------------------|-------------------------------|---------------------------|
| Á | V [˘] ^ Á [• äç ^ Á | V [˘] ^ Á ^ * äç ^ Á | V [˘] ^ Á äç ^ Á |
| Wäç Á | €Á | €Á | €Á |
| Ö } ^ i * ^ Á | €Á | €Á | €Á |
| Q f ! { äç } Á ^ & @ [[* ^ Á | €Á | €Á | €Á |

Á ä | ^ Á | K S | a - f ! • Á • d ^ • ~ | • Á - Ü ä ä [{ Ü ~ à • ^ c Á • ä * Á äç ^ Á c ä ^ • Á ä ç U Ô Á ä ç • ä Á

Á

| | | | |
|---|-------------------------------|-------------------------------|---------------------------|
| Á | V [˘] ^ Á [• äç ^ Á | V [˘] ^ Á ^ * äç ^ Á | V [˘] ^ Á äç ^ Á |
| | | | |
| | | | |
| | | | |

Ĥ ĒĀ æ\^cŪā ~|æ[!Á

Á

Y^ſ^ç^[] ſæ æ\^cŪā ~|æ[!Á Ā|[çā^ſ} Ā} çā[}{ ^} cŪ @!^ſ@ſ!æā * Á dæ^* ā•Á
&æ ſ^ſ•cāĒV@ſā ~|æ[!Á•^•Ā^æĒ [!Āç & Ā!æ•Ē} āſ @!æ ſā] |āā ſ^!•ā} Ā-Ā^æĀ
{ æ\ā

Î ÈÁÛ ðË -ÈÛæ] | ^Šãç ^ÁÛã ~ |æã } Á

Á

Î ÈÈÁÛ @æ] ^ÁÛæã Á

Á

ÁÛÁÛ | á | ÁÛ Á ^æ ~ | ^Áã dæ ^ * ç Á | ÷ | | { æ & È [| \ ã * Áæ ð • ó @ Á ç | ð ç Á ^ ç | } • Á @ Á
• dæ ^ ^ Á ^ } | | æ • Á Á [| ó] [~ * @ Á ^ Á ^ á Á Á | ÷ | | { æ & Á ^ æ ~ | ^ Á @ æ & } • ã | • Á @ Á
& } • ã ç } & Á Á @ Á ^ ç | } • Á ^ } | | æ á Á ^ Á dæ ^ È Á Á ã ç Á Á @ æ] ^ ç Á J I Á | ã ã ç Á ç | ã | ã ã
P ^ Á ç | á & á Á @ Á @ æ] ^ ÁÛæã Á @ æ Á Á ^ æ ~ | ^ Á | Á ç & | æã * Áã \ È æ ð • ç á Á ^ ç | } È V @ Á
Ù @ æ] ^ Á æã Á ~ ç Á ^ ç Á Á [| ç | ã Á ^ ç | } Á ã ~ • Áã \ È ^ Á æ Á ç á Á ç ã ^ á Á ^ ç } á æ á á Á ^ ç æã } Á
[- Á [| ç | ã Á ^ ç | } È

V @ Á @ æ] ^ ÁÛæã Á Á @ Á ç | | æ ^ Á ^ ç | } Á æ } ^ á Á Á ç &

Uc@!ÁÆ æ[] Ê^•c{ •Á@Á!ã^•Á} ÁÆ] |^ËYUT Á) áÁDÚÁ} á^!| ^!-!| { Á@Á æ\^Á
ã Á@Á~

P^i^E p a e ^ O e ^ . q ^ . l . A t A o l a a q * A ^ . c ^ { A } @ B A ^ . A p a e ^ O e ^ . A a o U U O A
[] c a a a e } A e A @ A } a ^ i | r q * A [a ^ | E a a A q a e | A a o U a a [{ A U ^ a ^ . c ^ A } @ B A ^ . l . A t A a
^ . c ^ { A } @ B A ^ { } [| ^ . A @ A U a a [{ A U ^ a ^ . c ^ A & @ a ~ ^ A a o U U O E] c a a ^ a A p a e ^ O e ^ . A e A @ A
~ } a ^ i | r q * A [a ^ | E A

p a e ^ O e ^ . A } A q a E A } ^ i * ^ A e a a q { | { a e } A ^ & @ [| ^ . A ^ & d | . A e | A e @ c ^ A [. a e ^ A
| ^ c | } . A ~ | a * A @ A . c a * A ^ i q a E A a e c | a A p a e ^ O e ^ . A a a q * A } A @ A | . A e A d & . A q A @ A
^ } ^ i * ^ A ^ & d | A e A a | { a a * A E | A e } ^ A a a E a q * A @ A c @ | A [A [A [A e ^ A e] ^ A a R A i r & p f r R & & 0
a e [c ^ A E V @ A U a a [{ A U ^ a ^

ĩ ěô [} & | ~ • ā } Á

Á

Q Á @ Á @

Ži áÚ@] ^ÉY a|ãä ÁZV@Á@] ^Áãä ÉVñ The journal of portfolio management ÁFÉÁFJJÍ DÁ JÉÌ ÈÁ
ŽGáÚ@] ^ÉY ZV@Á@] ^Áãä ÁÚ@] { |áÁM, ā^!•c ÉÁGFÍ ÉQj •HQ ^àÉ@] { |áÉá~ D, •@] ^ÉÉÉÉÉÁ
ŽÉáSimon, Phil. *Too Big to Ignore: The Business Case for Big Data*. Hoboken, New Jersey: John Wiley & Sons, Inc,
2013.Á
ŽFáÚ@] ^ÉY|ÁÚ| áÚ^áãä *Á áãã~ áÁã ä*•Á@] } *•Á| ^&@• ÉVñ Journal of Accounting Research ÁFJJÉDÁ
I €JÈ FÍ ÈÁ
ŽGáV• @ã äÚ| à^! ÉÚ^*!^•q } Á @ã \ æ ^Áã áÁ^!&ã } ÁãÁ@ Á@• [ÉVñ Journal of the Royal Statistical Society.
Series B (Methodological) ÁFJJÍ ÈÉ