כאן אדווח כל ניתוח סטטיסטי שאבצע.

## Trajectory Preprocessing

שלבים (מתוך Gallivan, J. P., & Chapman, C. S. (2014). Three-dimensional reach trajectories…) –

* מילוי דגימות חסרות – linear interpolation, או [inpaint\_nans](http://www.mathworks.com/matlabcentral/fileexchange/4551) של מטלאב.
* סינון רעש - low-pass Butterworth filter (dual pass, 8–12 Hz cutoff, 2nd order)

כרגע עושים lowpass 8Hz.

מתוך Using Functional Data Analysis\_v1\_april2011:

הסינון מבטל רעש, לכן אח"כ הוא עושה overfit למידע ללא חשש למידול רעש.

* קביעת ראשית צירים – עבור כל טרייל ההקלטה הראשונה היא ראשית הצירים.
* סיבוב מערכת צירים – עבור כל טרייל מסובבים אותה שתיהיה מאונכת למסך.

אצלנו הצירים מוגדרים כמאונכים למסך בתחילת הניסוי.

* מציאת רגעי תחילת וסוף התנועה – חילוץ וקטור מהירויות.

סינונו ב-low pass filter של 10-12Hz (אני לא עשיתי זאת).

תחילה - מציאת 4 דגימות ראשונות בהן המהירות גדולה מ-20mm/sec וגם

התאוצה הכוללת (ז"א לדעתי בין הראשונה לאחרונה) היא 20mm/s2.

סוף – מהירות קטנה מ-20mm/sec או מרחק מינימלי למסך, מה שמגיע ראשון.

* נרמול במרחב באמצעות b-spline – לא מימשתי בעצמי, השתמשתי בפונקציית normalize של קרייג.

~~נתאים b-spline למידע בכל ציר בנפרד (X,Y,Z) שיהיה מסדר 6 עם spline בכל דגימה.~~

~~b-spline היא פונקציה המחולקת לחלקים וכל חלק הוא פולינום.~~

~~Knot = המקומות בהם פונקציה אחת נגמרת ואחרת מתחילה~~

~~מספר knots הוא כמות ה-coefficients ועוד degree.~~

~~Degree = הסדר הגדול ביותר של הפולינומים שמרכיבים את ה-splines.~~

~~Order = degree + 1~~

~~B = basis function, כי הפונקציות הללו הן הבסיס לפונקציות spline אחרות, זה אומר~~

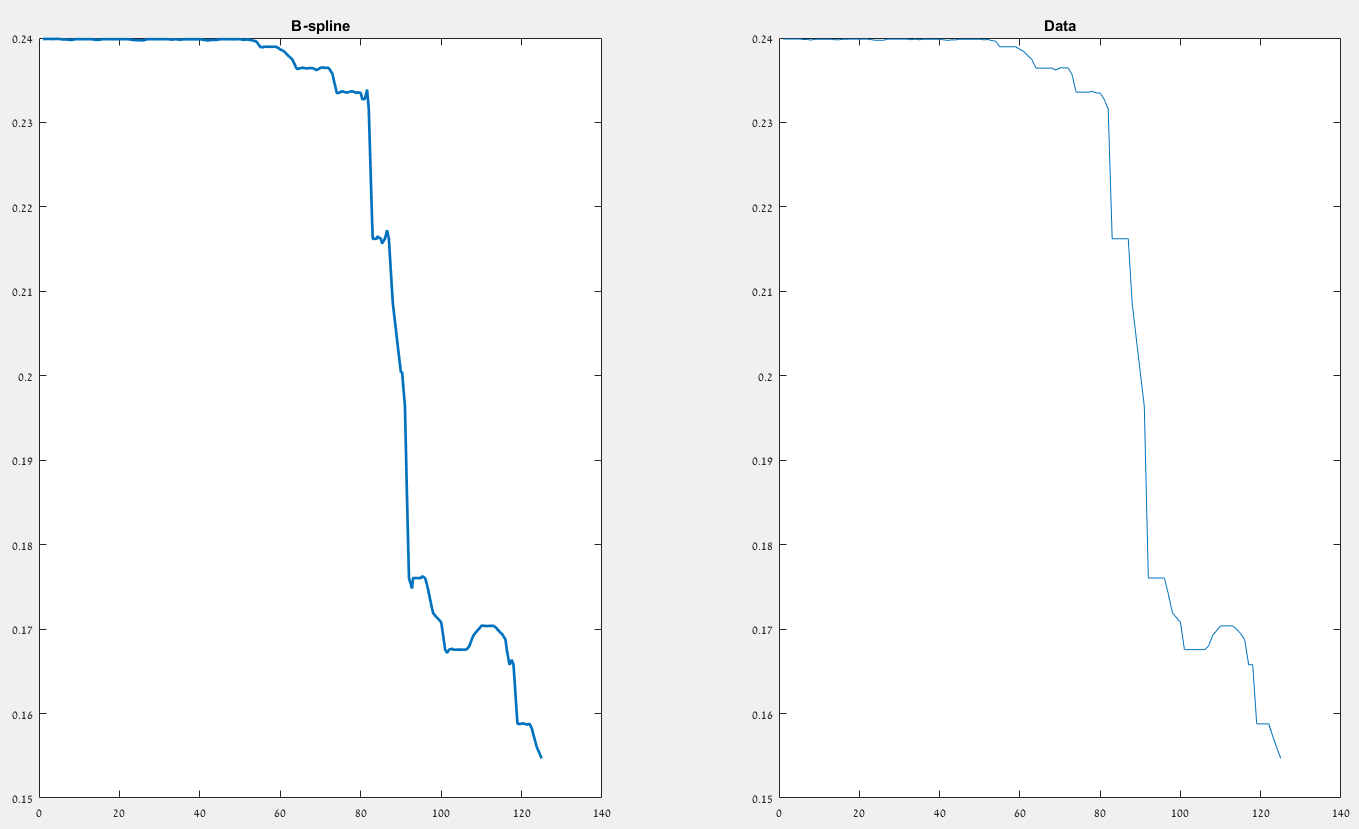
~~שאפשר להשתמש בבסיסיות כדי להרכיב את כל האחרות.~~

~~Cubic spline = פונקציית splines בה הפולינומים מדרגה 3.~~

~~C = continuity, מאפיין של spline שמציין כמה הוא חלק.~~

**~~% plotting: fnplt(b\_spline(1));~~**

**~~% getting values: fnval(b\_spline(1), 130.10)~~**

~~~~

~~~~

~~מתוך Using Functional Data Analysis\_v1\_april2011:~~

~~Bspline מסדר 6 שומר על המידע המקורי עד הנגזרת השלישית (=תאוצה).~~

**למה לנרמל?**

* + - * + סביבה משתנה בין נבדקים – סביר להניח שהמיקום של נקודת ההתחלה ושל המסך לא

יהיו זהים בין נבדקים. לפיכך נצטרך לנרמל את כולם לתבנית אחידה כדי שנוכל להשוות ביניהם.

* + - * + צמצום הבדלים בין נבדקים – האפקט שאנו מחפשים הוא תוך נבדקי (בין טריילים של

same ל-diff עבור כל נבדק). נרמול יצמצם הבדלים בזמני תגובה בין נבדקים, במסלול הושטת יד, במהירות הושטת יד.

* החלקה באמצעות roughness penalty –לא יישמתי בעצמתי, השתמשתי בשל קרייג. חלק מ-normalize לדעתי.

~~יישום roughness penalty על הנגזרת הרביעית, , שמיישמים ב-~~

~~within 0.00001 of the generalized cross validation estimate.~~

[~~אולי קשור~~](https://www.mathworks.com/help/curvefit/spaps.html)

[~~פונקציה להחלקה~~](https://www.mathworks.com/help/curvefit/csaps.html#d123e27998) ~~אבל לא ניתן להכניס את הנגזרת אותה רוצים להחליק.~~

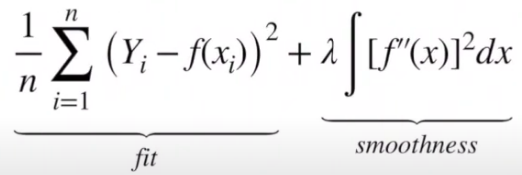
[~~מאמר על החלקה~~](https://sci-hub.se/https:/onlinelibrary.wiley.com/doi/abs/10.1002/0471667196.ess3138) ~~(roughness penalty)~~

[~~פונקציה ב-R~~](https://link.springer.com/content/pdf/10.1007%2F978-0-387-98185-7_5.pdf) ~~שעושה את ההחלקה המבוקשת~~

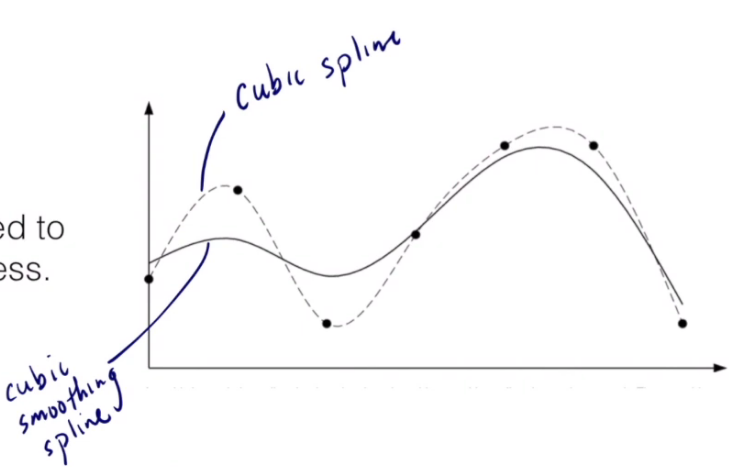
~~Roughness penalty = smoothing spline, שזה בעצם spline ששומר יותר~~

~~על החלקה ופחות על מעקב מדיוק אחרי נקודות המידע.~~

~~פונקציית הטעות אותה ממזערים לשם הערכת smoothing spline נראית כך:~~

~~~~

*~~= קובע האם ניתן משקל להתאמה למידע, או יותר להחלקה של הפונקציה.~~*

~~~~

* חילוץ נקודות מקבילות זו לזו במרחב – למה הוא מחלץ פעמיים נקודות? בהתחלה 2000 ואח"כ מתוכן 200?

מתוך Using Functional Data Analysis\_v1\_april2011:

מחלץ דגימות מציר Z (אליו מנרמלים) ברזולוציה גבוהה (2000 דגימות), ואז מוצא את הזמן של 200 מתוכן שהמרחק ביניהן על ציר Z שווה. לבסוף מחלץ ערכי X ו-Y בנקודות זמן אלו.

## Trial screening

נסנן את ה-trials **המקוריים** (לא אלו שעברו Preprocessing).

(בסינון נסיר traj אחד ולא את כל הטרייל, למשל נסיר רק את target\_x\_to אם הוא שגוי אבל לא את target\_x\_from, prime\_x\_to, prime\_x\_from)

נסיר את ה-trials הבאים:

* חסר הרבה data – פחות מ-100ms של מידע, או חסר יותר מ-100ms של מידע מהמסלול.
* מרחק הושטה קצר מדי – תנועה קצרה מ"מרחק בין מסך לנקודת התחלה" פחות 5 ס"מ.

נבדק על ה-preprocessed trials, לא המקוריים (כי זה המסלול שנכנס למיצוע).

* פספוס מטרה – הנבדק פספס את המטרה (לא נגע בתוך עיגול ברדיוס 3 ס"מ שמקיף את המטרה ומונח שטוח על המסך).

נסיר את הנבדקים ש:

* פחות מ-50% מהמידע שלהם עבר את הסינון (screening).
* פחת מ-100 טריילים בכל תנאי (אני החלטתי 100) עברו את הסינון.
* עשו הרבה טעויות סיווג (ייתכן שלא הבינו את המטלה או לא השקיעו בה). הסף יהיה מתחת ל-70% תשובות נכונות (קריטריון שאני הוספתי). לא מימשתי את זה

## Averaging

## בחינה ויזואלית של הנתונים

נמיין את המידע בצורות שונות ונראה האם יש אפקט? השונות גדולה מדי ולכן לא מומלץ למצע?

ניצור פונק':

* מיישרת צירים לנקודות על המסך – עושה טרנסופרמציה כך שראשית הצירים בין הנקודות וציר Z מאונך למסך.

מסתמכת על קליברציה בתחילת הניסוי.

* כיוון אחיד לכל הטריילים – לסובב לצד שמאל כל טרייל שהסיומת שלו חוצה לצד ימין את מחצית המסך.
* מפרידה בין same ל-diff
* מפרידה בין correct ל-wrong.
* סטיה בציר X – עבור טרייל בודד, ממוצע עבור קבוצת טריילים.

## ניתוח

* זמני תגובה congruent לעומת incongruent.
* הפרדת response time ל-reaction time (זמן תחילת תנועה) ו-movement time (משך תנועה).
* Maximum absolute deviation

מדד לאפקט, הנקודה הכי רחוקה מהקוו הישר שמחבר בין נקודת תחילת התנועה לנקודת סיום התנועה.



* מדדי ערך יחיד

נותנים ערך יחיד לכל טרייל ואותו משווים.

למשל הנקודה הכי רחוקה מהמרכז בכל טרייל.

* variability גבוהה מראה שאנשים מתלבטים, אז זה טוב שמצאנו זאת.
* אזור בין 2 מסלולים

מדד לאפקט, לחשב אזור בין same שמאלה ו-same ימינה, ואז להשוות לאותו הדבר ב-diff.



* מספר הפעמים שנבדק שינה את דעתו במסלול

מדד לאפקט.

Resulaj עשה את זה.



* שונות בתור מדד לאפקט.
* הפרש בין מסלולים

לחשב מסלול ממוצע לכל נבדק עבור כל תנאי, ואז לעשות חיסור לכל נבדק בין 2 התנאים.

כעת למצע את ההפרש הזה בין נבדקים ולצייר אותו לאורך הזמן. להוסיף STD, ואז מקבלים מבחן t לאורך זמן.



* ניתוח זווית (כיוון) בכל נקודה בזמן.



* בטרייל מסוים, מיקום בכל נקודה זמן (נוכל לראות אם הנבדק מתמהמה ב-incongruent).

לנתח כל ציר בנפרד?

נרמול: נחלק כל נקודה בזמן, באורך הכולל של הטרייל. נקבל מיקום בכל נקודת זמן מ-0 ועד 1.

חסרון: מוחק הבדלים בזמני תגובה.

* בטרייל מסוים, סטיה מציר X לאורך הזמן (נוכל לראות אם הנבדק סוטה מאוחר יותר מהאמצע ב-incongruent).

נרמול: נחלק כל נקודה בזמן, באורך הכולל של הטרייל. נקבל מיקום בכל נקודת זמן מ-0 ועד 1.

נחלק כל נקודה ב-X במרחק בין נקודת התחלה לנקודה סופית.

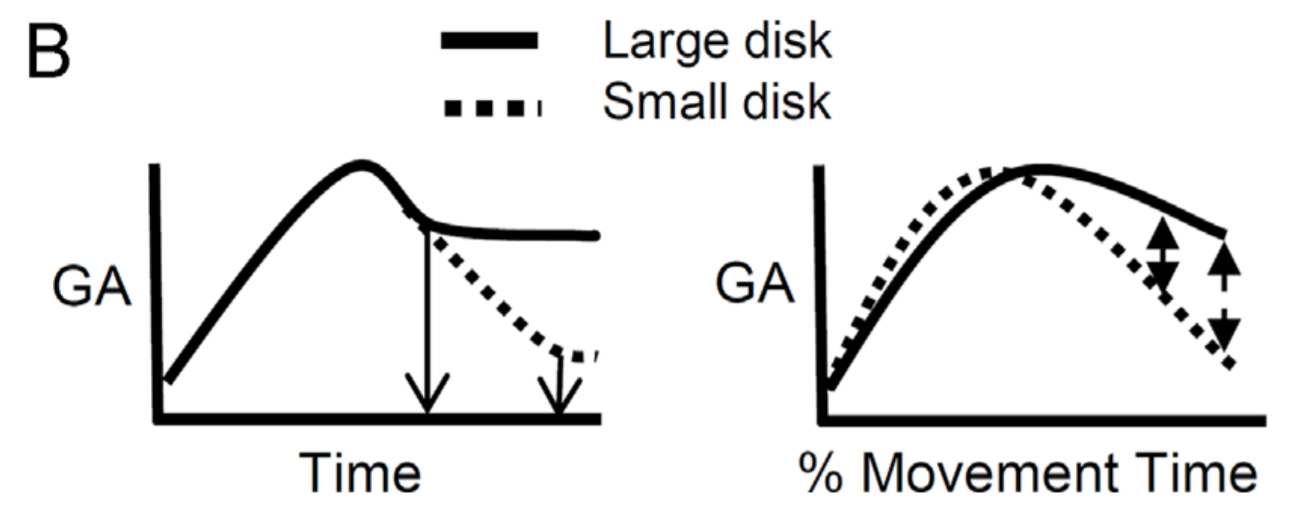
חסרון: מוחק הבדלים בזמני תגובה.

מוחק הבדלים במסלול עצמו (למשל מצבים של סטיה קטנה יותר בציר X עבור incongruent)

* האם לנרמל לפי טרייל או לפי ממוצע של הנבדק?
* לחפש מידע לגבי נרמול:
  + ~~Dotan - Track it to crack it~~
  + ~~Dotan, D. et al. (2018) On-line confidence monitoring during decision making~~
    - חישוב מהירות – עשו נגזרת למסלול בציר מסוים כדי לקבל את המהירות בו.
    - לפני נגזרת עשו gaussian smoothing (σ=20ms).
  + ~~Cressman, E.K. et al. (2007) On-line control of pointing is modified by unseen visual shapes. Conscious. Cogn. 16, 265–275 – חישבו מסלול ממוצע (ממוצע בכל נק' בזמן בין כל הטריילים).~~
  + בעיות בנרמול בזמן
    - ~~Whitwell 2013 - Grasping without vision time normalizing grip aperture profiles yields spurious grip scaling to target size~~

**לסיכום –** כשמשך התנועה משתנה במתאם עם המשתנה הבלתי תלוי, ואז מנרמלים לפי זמן, נצטרך לשים לב לכך שנקודות זהות בזמן המנורמל מייצגות נקודות שונות בזמן האמיתי.

למשל אם ב-2 התנאים יש מסלול זהה פרט לחלק קטן נוסף בסוף המסלול של תנאי 2, אחרי שננרמל יראה לנו שיש הבדל בנקודת ההתחלה!

(מטלת אחיזה בלי לראות את המטרה. גרף שמאלי לפני נרמול גרף ימני אחרי, בו נדמה שאחיזה בדיסקים קטנים

מתחילה מוקדם יותר).

* + - מתי להימנע מנרמול בזמן: כשמתמקדים בחלק מוקדם של המסלול:

~~. Dotan, D. and Dehaene, S. (2016) On the origins of logarithmic number-to-position mapping. Psychol. Rev. 123, 637–666 לא קראתי הכל, אבל לא ראיתי שם משהו שקשור לנרמול.~~

~~4. Dotan, D. and Dehaene, S. (2013) How do we convert a number into a finger trajectory? Cognition 129, 512–529~~

~~לא קראתי הכל, אבל לא ראיתי שם משהו שקשור לנרמול.~~

### Linear mixed model

#### Reach area

נראה שה-interpolation של ה-trajectories לא עובדת טוב מה שגורם למסלולים מעוותים.

אדום = הושטה שמאלה (כן זה הפוך), ירוק = הושטה ימינה.

Chart, line chart

Description automatically generated

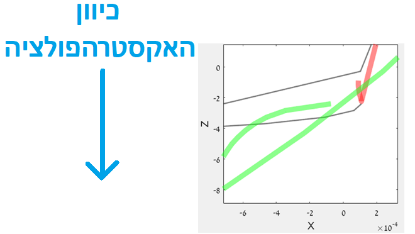
זה יוצר ערכים קיצוניים ב-reach area.

Chart, waterfall chart

Description automatically generated

מבט מקרוב מגלה שהמסלול האדום מתחיל בנקודה מאוחרת על ציר מהמסלול הירוק, משמע שצריך לעשות לו אקסטרהפולציה כדי לקבל את ערכי X של המסלול האדום בכל אחת מנקודות ה-Z של המסלול הירוק.

הבעיה היא שהמסלול האדום לא פונה לכיוון אליו רוצים לעשות אקסטרה פולציה, ולכן ה-fit צריך להתאים ערכים קיצוניים כדי לנסות לפצות על כך.



כדי לפתור זאת נציב סף על ציר Z, ונחשב את השטח רק מעבר לסף זה.

כדי להימנע מביקורת עתידית, ניתן להכניס רעש לתחילת המסלול ולהראות ששינוי הסף אינו יוצר אפקט בשל הרעש.

* Reach area
  + נחפש את הנבדק שמספר הטריילים שיש לו בתנאי בודד הוא הקטן ביותר מכל הנבדקים, נקרא לו X.
  + ניצור התפלגות "מזויפת" של ה-trajectory הממוצע שלו, ע"י דגימה (עם החזרה) של X טריילים ומיצוע שלהם, 1000 פעמים.

נחשב את השטח בין בין ה-trajectories הממוצעים.

נתייחס לכל ממוצע כזה כתצפית אחת במודל שלנו.

נעשה זאת עבור כל נבדק בכל תנאי.

* + נבנה mixed model שמנסה לנבא את ה-reach area, המנבאים יהיו: מספר נבדק (random), תנאי.
* MAD
* Trajectory
  + נבנה mixed model שמנסה לנבא את המיקום על ציר X, המנבאים יהיו: מספר נבדק (random), תנאי, צד.
  + נכין מודל כזה לכל נקודה לאורך ציר Z.
  + נערוך FDR כדי לתקן השוואות מרובות?