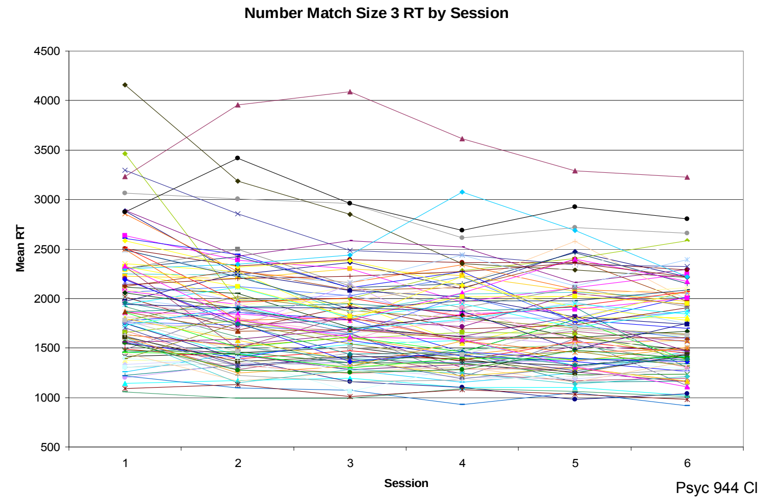
Polynomial Models of Change

- Today's topics:
  - Big Picture of Modeling Change
  - Modeling Means with Polynomial Fixed Effects
  - Modeling Variances with Polynomial Random Effects
  - Interpreting Random Effects Variances using Confidence Intervals
  - Wrapping up…

Example Data Individual Observed Trajectories (*N* = 101, *n* = 6)



Describing Within-Person Change: ***Model for the Means (Fixed Effects)***

- What kind of change occurs on average?
  - **What is the most appropriate metric of time?**
    - Time in study *(with predictors for BP differences in 'time')*?
    - Time since birth *(age)*? Time to event *(time since diagnosis)*?
    - Measurement occasions need not be the same across persons or equally spaced (code time as exactly as possible)
  - **What kind of population model generated the observed trajectories?**
    - Linear or nonlinear? Continuous or discontinuous? One process or multiple processes?

Big Picture Modeling Framework: *Choices for Each Side of Model*

- **Model for the Means: Predicted means across time**
  - "**Empty**" refers to model for the means with no predictors (just fixed intercept for **grand mean** outcome over time)
  - "**Saturated**" refers to model for means with all possible means estimated (#parameters = #occasions) → **THIS IS ANOVA**
    - Is a DESCRIPTION of means, not a parsimonious MODEL

Parsimony

__**Empty Model:**__ Predicts NO change over time
1 Fixed Effect

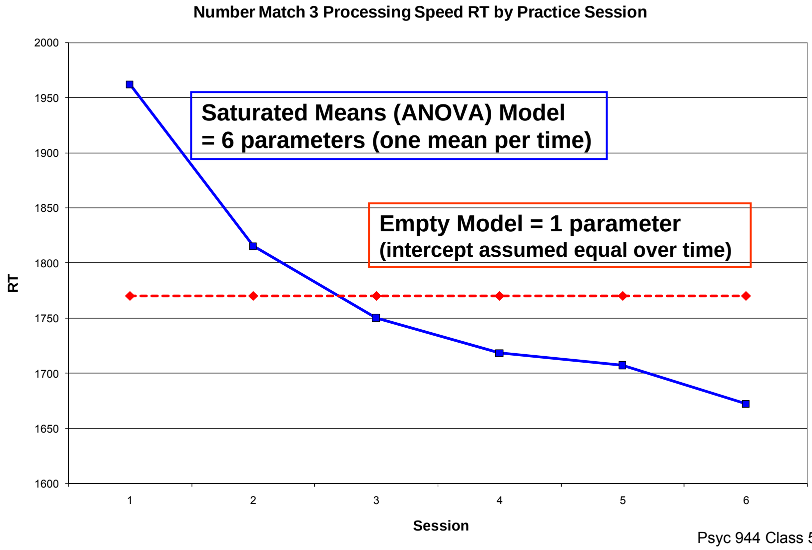
Most useful model: likely somewhere in between!

__**Saturated Means:**__ Reproduces mean at each occasion
 # Fixed Effects = # Occasions

Good fit

Name… that… Trajectory!

Baseline Fixed Effects Models



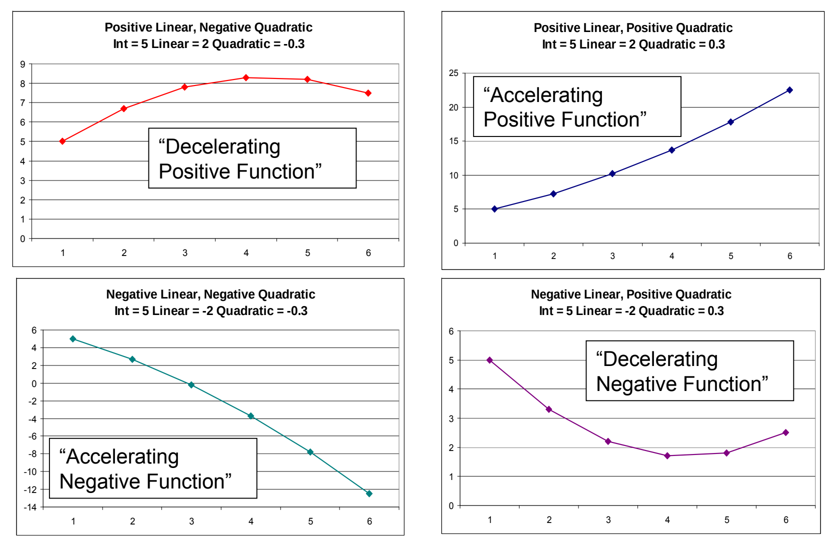
Name that Trajectory with Polynomial Fixed Effects of Time

- Represent **mean** patterns of change with polynomial fixed effects:
  - Linear → *constant* amount of change (up or down)
  - Quadratic → *change* in linear rate of change (acceleration/deceleration)
  - Cubic → *change* in acceleration/deceleration in linear rate of change
  - Terms work together to describe curved trajectories
- Can have fixed effect polynomials UP TO: # occasions – 1
  - 3 occasions = 2nd order = Fixed Quadratic or less
  - 4 occasions = 3rd order = Fixed Cubic or less
- Interpretable polynomials past cubic are rarely seen in practice – just because you *could* estimate higher orders doesn't mean you *should*!

Lower-Order *Fixed* Effects Are Conditional on Higher-Order *Fixed* Effects

- You can put the intercept (set time=0) anywhere you want, and you will get the *same model fit and predicted values*.
- Different centerings of time will yield models with different estimates and interpretations for the *lower-order terms*, however:
  - ***Fixed Intercept Only?***
    - Fixed Intercept = mean of Y *for any time point (= grand mean)*
  - ***Add Fixed Linear?***
    - Fixed Intercept = **now** mean of Y *when time=0*
    - Fixed Linear = mean linear rate of change *across all occasions*
  - ***Add Fixed Quadratic?***
    - Fixed Intercept = still mean of Y *when time=0*
    - Fixed Linear = **now** mean linear rate of change *when time=0*
    - Fixed Quadratic = mean acceleration/decel. of change *across all occasions*

Examples of Fixed Quadratic Models

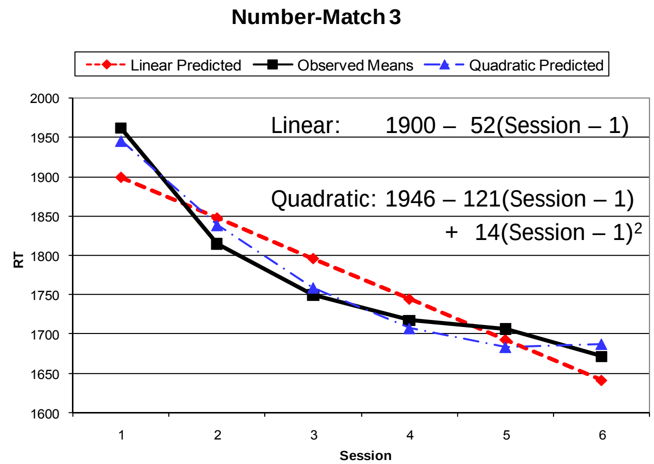


Interpreting Quadratic Coefficients

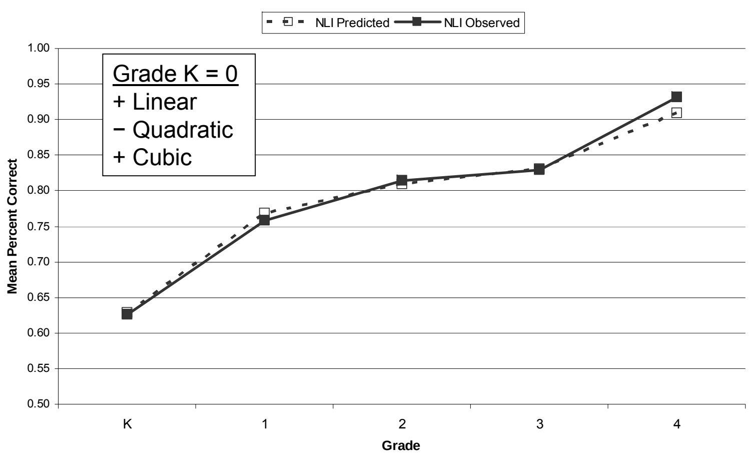
A Quadratic effect is a two-way interaction: time*time

- Fixed quadratic coefficient = "**rate of acceleration/deceleration**"
- To interpret it as *how the linear trend changes per unit time*, however, **you must multiply the quadratic coefficient by 2**
- If linear trend = 5 at time 0, with quadratic = .5?
  - Instantaneous linear rate of Δ at time 0 = 5, at time 1 = 6, at time 2 = 7…
- The "twice" part comes from taking the derivatives of the function:
  - Original function → y = 10 + 5(time) + .5(time)^2
  - Linear rate of Δ = 1st derivative → y' = __ + 5 + **1**(time)
  - Slope of rate of Δ = 2nd derivative → y'' = __ + __ + **1**
- Put another way: Because time is interacting with itself, there is no second 'main effect' in the model. So the quadratic effect gets applied to time twice, given that there is only one main effect of time.

Example Data:
Observed vs. Model-Predicted Means



Real-Life Example: Fixed Cubic Model
Growth of Grammar Skills in Kids with Language Impairments



Back to the Big Picture: Modeling Means and Variances

- We will have two tasks in describing longitudinal data:

1. Choose a Model for the Means

- What kind of change in the outcome do we have on average?
- What kind of and how many parameters do we need to represent that change as parsimoniously but accurately as possible?

2. Choose a Model for the Variances

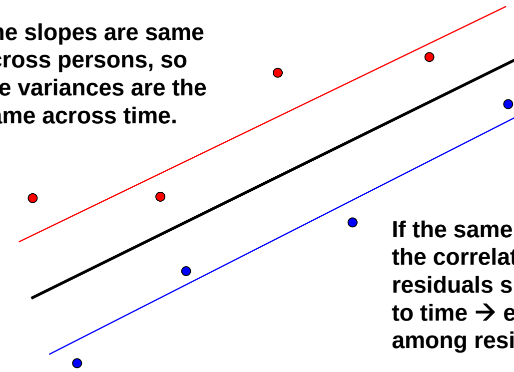
- What kind of pattern do the variances and covariances of the outcome show over time?
- What kind of and how many parameters do we need to represent that pattern as parsimoniously but accurately as possible?

Describing Within-Person Change: *Model for the Variances (Random Effects)*

- ***From a substantive perspective:* Are there individual differences in change?**
  - Individual differences in level?
    - At what time point are individual differences in level important for your hypotheses (beginning, middle, end)?
  - Individual differences in magnitude of change?
    - Each aspect of change (e.g., linear change, quadratic change) can potentially exhibit individual differences (data permitting)
- ***From a statistical perspective:* What kind of pattern do the variances and covariances exhibit over time?**
  - Do variances increase or decrease over time?
  - Are covariances for closer occasions more related?

Choices in Modeling Variances: Random Intercept Only (Compound Symmetry)

The slopes are same across persons, so the variances are the same across time.

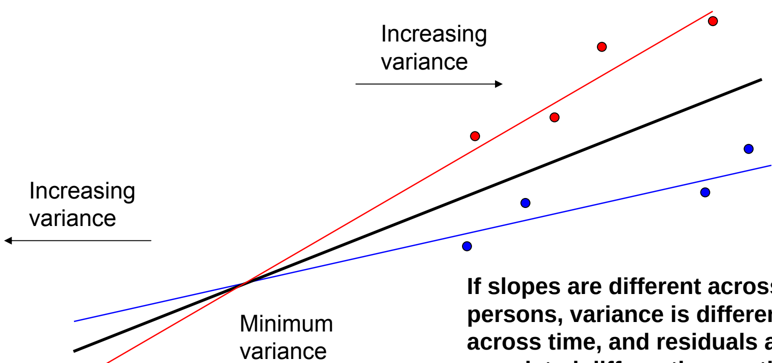


$$\begin{bmatrix} \sigma_e^2+\tau_{u_0}^2 & & & \\ \tau_{u_0}^2 & \sigma_e^2+\tau_{u_0}^2 & & \\ \tau_{u_0}^2 & \tau_{u_0}^2 & \sigma_e^2+\tau_{u_0}^2 & \\ \tau_{u_0}^2 & \tau_{u_0}^2 & \tau_{u_0}^2 & \sigma_e^2+\tau_{u_0}^2 \end{bmatrix}$$

If the same slope fits all, then the correlations among residuals should not be related to time → equal correlation among residuals across time.

If slopes are the same across people, then people differ from each other systematically in only 1 way (i.e., their U_i level) → THIS IS COMPOUND SYMMETRY.

Choices in Modeling Variances: Random Intercepts and Slopes Model



If slopes are different across persons, variance is different across time, and residuals are correlated differently over time.

If slopes are different across people, then people differ from each other systematically in 2 ways (U_{0i} and U_{1i}) → this implies compound symmetry will NOT hold.

Big Picture Modeling Framework: Choices in Models for the Variances

Compound Symmetry (CS)

$$\begin{bmatrix} \sigma_e^2+\tau_{u_0}^2 & & & \\ \tau_{u_0}^2 & \sigma_e^2+\tau_{u_0}^2 & & \\ \tau_{u_0}^2 & \tau_{u_0}^2 & \sigma_e^2+\tau_{u_0}^2 & \\ \tau_{u_0}^2 & \tau_{u_0}^2 & \tau_{u_0}^2 & \sigma_e^2+\tau_{u_0}^2 \end{bmatrix}$$

Most useful model: likely somewhere in between!

Unstructured (UN)

$$\begin{bmatrix} \sigma_{11}^2 & & & \\ \sigma_{21} & \sigma_{22}^2 & & \\ \sigma_{31} & \sigma_{32} & \sigma_{33}^2 & \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_{44}^2 \end{bmatrix}$$

Parsimony ← → Good fit

Univ. WP ANOVA — **NAME ...THAT ... STRUCTURE!** — Multiv. WP ANOVA

CS and UN are just 2 of the many, many options available for modeling the variances in MLM.
There are two main families: (1) **random effects models** and (2) alternative covariance structure models.

Polynomial Fixed vs. Random Effects

- Polynomial **fixed effects** combine to describe linear or non-linear **sample mean trajectories** over time (**up to #occasions – 1**):
  - Fixed Intercept = **Mean** level (where time=0)
  - Fixed Linear = **Mean** linear rate of change (where time=0)
  - Fixed Quadratic = **Mean** acceleration/deceleration in linear rate of change

- Polynomial **random effects** (individual deviations from the sample mean for each change parameter) describe **individual differences** in those change parameters (**up to #occasions – 2**):
  - Random Intercept = **BP Variance** (individual diffs) in intercept level
  - Random Linear = **BP Variance** (individual diffs) in linear rate of change
  - Random Quadratic = **BP Variance** (individual diffs) in acceleration or deceleration of linear rate of change

Random Effects Allowed by #Occasions

	V Matrix		**G Matrix**	**R Matrix**
2 time points = **3** pieces of information	$\begin{bmatrix}\sigma_{11}^2 & \\ \sigma_{21} & \sigma_{22}^2\end{bmatrix}$	→	$\begin{pmatrix}\tau_{U_0}^2\end{pmatrix}$ Random Intercept only	$\left(\sigma_e^2\right)$ Assume residuals are constant and uncorrelated across all times and people
3 time points = **6** pieces of information	$\begin{bmatrix}\sigma_{11}^2 & & \\ \sigma_{21} & \sigma_{22}^2 & \\ \sigma_{31} & \sigma_{32} & \sigma_{33}^2\end{bmatrix}$	→	$\begin{pmatrix}\tau_{U_0}^2 & \\ \tau_{U_{01}} & \tau_{U_1}^2\end{pmatrix}$ Up to 1 Random slope	$\left(\sigma_e^2\right)$ Assume residuals are constant and uncorrelated across all times and people
4 time points = **10** pieces of information	$\begin{bmatrix}\sigma_{11}^2 & & & \\ \sigma_{21} & \sigma_{22}^2 & & \\ \sigma_{31} & \sigma_{32} & \sigma_{33}^2 & \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_{44}^2\end{bmatrix}$	→	$\begin{pmatrix}\tau_{U_0}^2 & & \\ \tau_{U_{01}} & \tau_{U_1}^2 & \\ \tau_{U_{02}} & \tau_{U_{12}} & \tau_{U_2}^2\end{pmatrix}$ Up to 2 Random slopes	$\left(\sigma_e^2\right)$ Assume residuals are constant and uncorrelated across all times and people

Predicted Matrices from Random Effects Models

- **Random linear model?** Variance has a **quadratic** dependence on time
  - Variance will be at a minimum when time = $-\text{Cov}(U_0,U_1)/\text{Var}(U_1)$, and will increase parabolically and symmetrically over time
  - **Predicted variance** at each occasion:

$$\begin{aligned}\text{Var}(Y_t) &= \text{Var}(e_t + U_0 + U_1{*}\text{time}_t)\\ &= \text{Var}(e_t) + \text{Var}(U_0) + \text{Var}(U_1)(\text{time}_t^2) + 2\text{Cov}(U_0,U_1)(\text{time}_t)\end{aligned}$$

  - **Predicted covariance** between occasions *a* and *b*:

$$\text{Cov}(Y_{ta},Y_{tb}) = \text{Var}(U_0) + \text{Cov}(U_0,U_1)(\text{time}_a + \text{time}_b) + \text{Var}(U_{1i})(\text{time}_a)(\text{time}_b)$$

- **Random quadratic model?** Variance has a **cubic** dependence on time

- *The point of the story – random effects are a way of allowing the variances and covariances to differ over time in specific, time dependent patterns that arise due to change over time.*

Lower-Order *Random* Effects Are Conditional on Higher-Order *Random* Effects

- As with fixed effects, different centerings will yield equivalent models with different estimates and interpretations for the *lower-order terms*:
  - ***Random Intercept Only?***
    - Random Intercept = individual differences *for any time point* in mean Y (= variance in grand mean because individual lines are parallel)
  - ***Add Random Linear?***
    - Random Intercept = **now** individual differences *when time=0* in mean Y
    - Random Linear = individual differences *across all occasions* in linear rate of change *(would be the same if time was centered elsewhere)*
  - ***Add Random Quadratic?***
    - Random Intercept = still individual differences *when time=0* in mean Y
    - Random Linear = **now** individual differences *when time=0* in linear rate of change *(would be different if time was centered elsewhere)*
    - Random Quadratic = individual differences *across all occasions* in accel. or decel. of change *(would be the same if time was centered elsewhere)*

Polynomial Fixed vs. Random Effects

- On the same side of the model (means or variances side), lower-order effects stay in EVEN IF NONSIGNIFICANT (for correct interpretation)
  - E.g., Significant fixed quadratic? Keep the fixed linear, no matter what*
  - E.g., Significant random quadratic? Keep the random linear
- Also remember – you can have a significant random effect EVEN IF the corresponding fixed effect is not significant:
  - Fixed effect stays in the model anyway (clearer interpretation)
  - E.g., Fixed linear not significant, but random linear is significant? → No change *on average*, but significant individual differences in change
- Language: A random effect supersedes a fixed effect:
  - If <u>Fixed</u> = intercept, linear, quad; <u>Random</u> = intercept, linear, quad?
    - Call it a "Random quadratic model" (implies everything beneath those terms)
  - If <u>Fixed</u> = intercept, linear, quad; <u>Random</u> = intercept, linear?
    - Call it a "Fixed quadratic, random linear model" (distinguishes no random quad)

Recommended Sequence for Testing Fixed and Random Effects of Time

<u>Build up fixed and random effects simultaneously:</u>

1. Random Intercept only (Empty Model; calculate ICC)
2. Fixed Linear, Random Intercept
3. Random Linear
4. Fixed Quadratic, Random Linear
5. Random Quadratic
6. …….

<u>*** Note: USE **REML** for all models, but make sure to:</u>
→ Test significance of new **fixed** effects by **p-values**
→ Test significance of new **random** effects by **deviance differences**
→ Also examine AIC and BIC to see if new term(s) are worth it

Interpreting Random Effects Variances
(see Snijders & Bosker p. 48-50)

- Deviance difference tests will tell you whether or not a given random effect is significant…
  - In other words, if there are significant individual differences in that random effect (i.e., if everyone needs their own intercept or slope)
- But variance components aren't directly meaningful…
  - "I have a random intercept variance of 273,306. It's significant, but what does that actually *mean*? Is 273,306 a lot?"
- Interpret the magnitude of variance components by constructing *95% Random Effects Confidence Intervals*
  - Predicted range for that fixed effect under which 95% of the sample is predicted to fall, using the original metric of the variable

Interpreting Random Effects Variances

- **Random Effects 95% CI = fixed effect ± (1.96 * SQRT(variance))**
  → Predicted values assuming a normal, symmetric distribution!!
  → If yours isn't normal, these can go out of bounds of the possible data
- Example: fixed intercept = 1946, random intercept variance = 273,306
  = 1946 ± (1.96 * SQRT(273,306)) → 1946 ± 1046 → 900 to 2992
  = 95% of my sample is predicted to have an individual intercept between 900 and 2992
  - Can do the same for each random effect (e.g., intercept, slopes…)
- NOT the same as a typical CI around the point estimate
  - <u>Point estimate CI</u> tells you what the **fixed effect** is predicted to be 95% of the time (describes uncertainty of point estimate for mean effect)
  - <u>Random effects CI</u> tells you the range of the fixed effect for 95% of the **individuals** in your sample (describes variation across people)

What about the covariances/correlations between random effects?

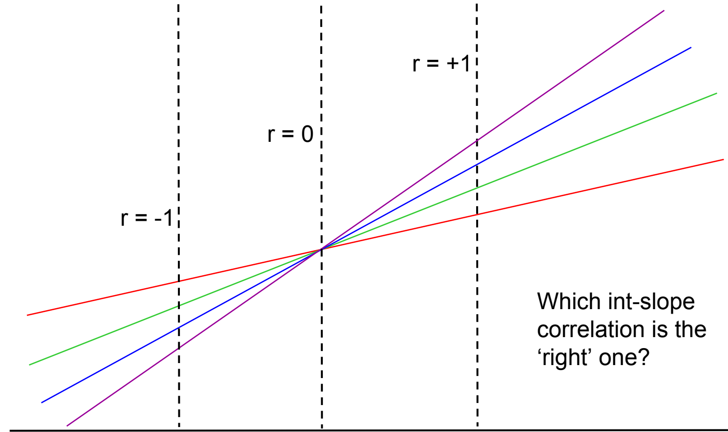
Should covariances always be **included**?

- For longitudinal models, **YES!** Include by specifying:
  - TYPE=UN on SAS RANDOM statement (default is TYPE=VC)
  - COVTYPE(UN) on SPSS RANDOM statement (default is TYPE=VC)

Should covariances always be **interpreted**?

- For longitudinal models, **not necessarily**
- Correlation between random effects will depend on centering of time (as will variance of lower-order random effects)
- Intercept-Slope Correlation is NOT interpretable unless:
  - Time 1 really is the beginning of the process under study
  - Intercept is coded as the mean across time points
    - Even then, it's interpretable only within the range of time measured

Correlation Between Intercept & Slope



!! Nonparallel lines will eventually cross.

Session Centered at 1:

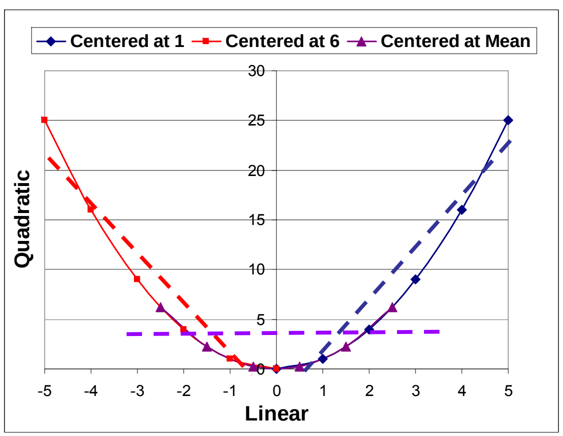
Session	Linear	Quadratic
1	0	0
2	1	1
3	2	4
4	3	9
5	4	16
6	5	25

Session Centered at 6:

Session	Linear	Quadratic
1	-5	25
2	-4	16
3	-3	9
4	-2	4
5	-1	1
6	0	0

Session Centered at Mean:

Session	Linear	Quadratic
1	-2.5	6.25
2	-1.5	2.25
3	-0.5	0.25
4	0.5	0.25
5	1.5	2.25
6	2.5	6.25



Correlations among growth terms can be induced by centering time near the start or near the end.

Correlations are *most* interpretable when centering time at its mean.

Wrapping up…

- Modeling within-person change involves specifying both sides of the model
  - Fixed effects in model for the means:
    - What kind of change am I observing on average?
    - What kind of trajectory will reproduce those means?
  - Random effects (and residuals) in model for the variances:
    - What kind of individual differences in change am I observing?
    - How many random effects do I need to reproduce the observed matrix of variances and covariances over time?
- One option: Polynomial models (linear, quadratic, cubic)
  - Terms work together to describe non-linear trajectories
  - Careful with the covariances among random effects, though
- Coming later: Piecewise models of change…