

Interakcje między zmiennymi ciągłymi w regresji wielokrotnej

Bartosz Maćkiewicz

W tej części kursu zajmiemy się *interakcjami* między zmiennymi w regresji wielokrotnej. Temat podzielimy na dwie części. Pierwsza z nich poświęcona będzie zmiennym ciągłym, druga zaś nominalnym. Problematyka interakcji wygląda dość podobnie dla obu rodzajów zmiennych, niemniej jest kilka istotnych różnic, które należy omówić. Dziś będzie trudniejsza część, czyli zmienne ciągłe.

Na dzisiejszych zajęciach korzystać będziemy ze znanych już Państwu pakietów `pander`, `knitr`, `car` oraz `carData`. Pracować będziemy na zbiorze danych, którego dotyczyło zadanie domowe z poprzedniego, czyli na zbiorze `Salaries` dostępnym w pakiecie `carData`, w którym znajdują się dane o wynagrodzeniach pracowników uniwersyteckich w Stanach Zjednoczonych. Na początek wczytamy odpowiednie pakiety.

```
library(pander)
library(knitr)
library(carData)
library(car)
```

Pojęcie interakcji

Pojęcie interakcji często wprowadza konfuzję u tych, którzy go jeszcze nie znają. W tej sekcji skoncentrujemy się więc na próbie intuicyjnego uchwycenia tego pojęcia. W tym celu posłużę się przykładem z książki “Applied Multiple Regression” (Cohen, Cohen, West, Aiken 2003), którą mają Państwo dostępną w materiałach dla zalogowanych.

“Wiele teorii w naukach społecznych stawia hipotezę, że dwie (lub więcej) ciągle zmienne wchodzi w interakcję. Rozważmy jako przykład problem tego jak umiejętności (X) oraz motywacja (Z) wpływają na osiągnięcia na studiach wyższych [*graduate school* nie ma dobrego odpowiednika, ale w tym wypadku nie ma to żadnego znaczenia - BM]. Pierwsza możliwość jest taka, że ich efekty się ze sobą kumulują. Połączony wpływ umiejętności i motywacji na osiągnięcia będzie równy sumie odrębnych efektów; brak interakcji między X i Z. Można powiedzieć, że

całość równa się sumie części. Drugą alternatywą jest to, że umiejętności i motywacja wchodzi w synergistyczną interakcję, tak że studenci z zarówno wysokimi umiejętnościami jak i wysoką motywacją osiągają znacznie więcej na studiach niż przewidywalibyśmy na podstawie efektów prostej sumy efektów umiejętności i motywacji. Studenci, którzy brylują w obu tych aspektach stają się „gwiazdami”; powiedzielibyśmy w tym przypadku, że całość jest większa, niż jej części. Trzecia alternatywa jest taka, że umiejętności oraz motywacja kompensują się wzajemnie. Dla tych studentów, którzy mają wyjątkowo wysoki poziom umiejętności, motywacja nie jest tak istotna dla osiągnięcia sukcesu, podczas gdy dla studentów z najwyższą motywacją, czyste umiejętności nie mają znaczenia. Tutaj powiedzielibyśmy, że całość stanowi mniej niż suma części, występuje częściowy kompromis (*trade-off*) między umiejętnościami i motywacją jeśli chodzi o przewidywanie osiągnięć akademickich. Druga i trzecia alternatywa stanowi przykład interakcji między predyktorami, to znaczy łącznego efektu predyktorów, który różni się od sumy ich odrębnych efektów.” (s. 255, tłumaczenie niestety własne)

Widzimy zatem, że badanie interakcji może być bardzo przydatną techniką dla każdego badacza, w tym oczywiście także kognitywisty :). Spróbujmy więc w praktyce zademonstrować, jak wygląda w R badanie interakcji i na co trzeba uważać, tworząc modele z interakcjami.

Regresja bez interakcji

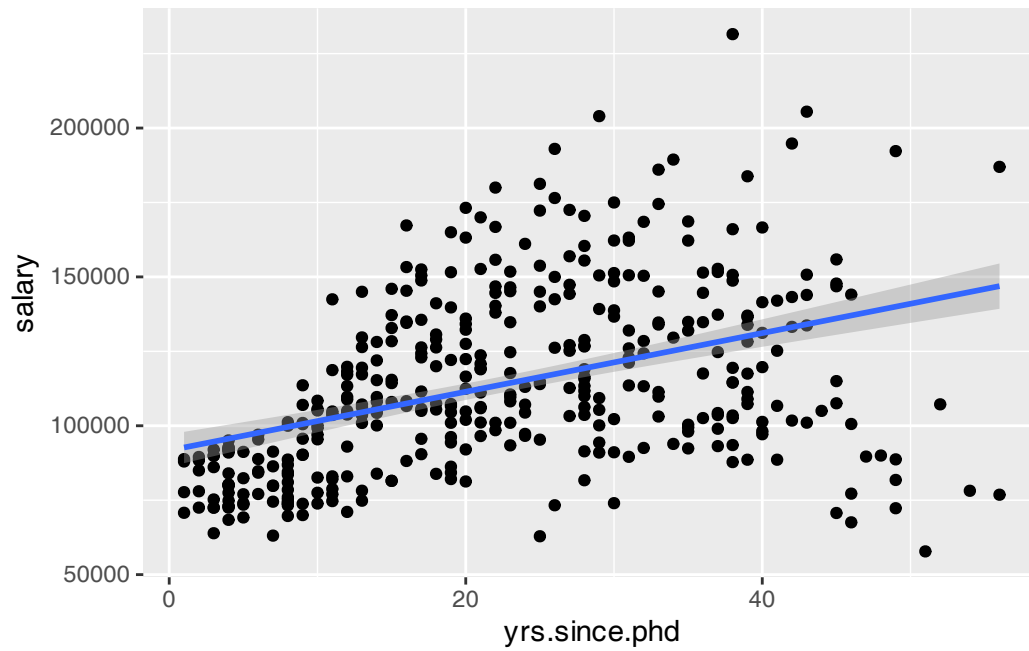
Zbiór danych, na którym będziemy pracować, dotyczy wynagrodzenia pracowników akademickich. W kolumnie `salary` znajduje się wysokość rocznego wynagrodzenia. Kolumny `yrs.service` oraz `yrs.since.phd` zawierają informacje odpowiednio o stażu w danej instytucji oraz liczbie lat, która upłynęła od uzyskania stopnia doktora.

Na początek zobaczmy, jak wyglądają relacje między zmiennymi `yrs.service` i `yrs.since.phd` oraz `salary`, kiedy weźmiemy je z osobna. W tym przypadku dobrym sposobem będzie narysowanie odpowiednich wykresów rozrzutu z nałożonymi na nie liniami regresji.

Dla wynagrodzenia i lat, które upłynęły od uzyskania stopnia doktora:

```
library(ggplot2)
ggplot(Salaries, aes(x=yrs.since.phd, y=salary)) +
  geom_point() +
  geom_smooth(method=lm)
```

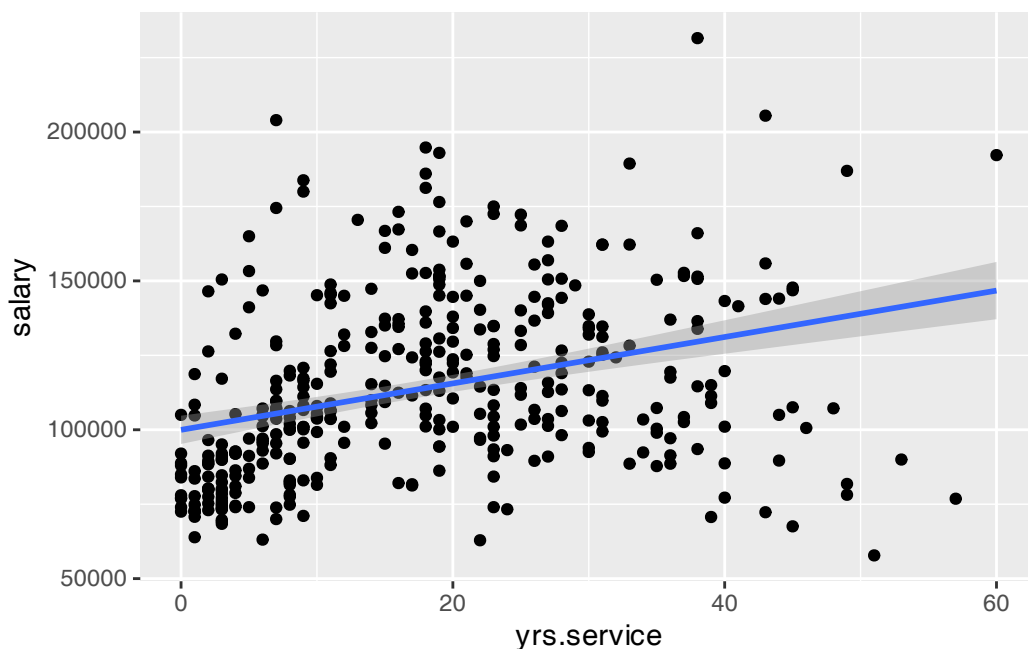
```
`geom_smooth()` using formula = 'y ~ x'
```



Dla wynagrodzenia i stażu:

```
ggplot(Salaries, aes(x=yrs.service, y=salary)) +  
  geom_point() +  
  geom_smooth(method=lm)
```

`geom\_smooth()` using formula = 'y ~ x'



Widzimy więc, że jeśli rozważamy relacje między każdą zmienną objaśniającą z osobna i zmienną objaśnianą, to są one pozytywne - im wyższy ktoś ma staż, tym więcej zarabia na akademii, im więcej upłynęło od doktoratu, tym zarobki są większe.

Teraz stwórzmy sobie teraz prosty model, w którym predyktorami będą zmienne `yrs.since.phd` (liczba lat, które upłynęły od uzyskania doktoratu) oraz `yrs.service` (staż pracy), a wartością, którą chcemy przewidywać, będzie oczywiście wynagrodzenie (wyrażone, jak mam nadzieję, w dolarach rocznie).

```
fit <- lm(salary ~ yrs.since.phd + yrs.service, data = Salaries)
pander(summary(fit))
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	89912	2844	31.62	3.811e-110
yrs.since.phd	1563	256.8	6.086	2.754e-09
yrs.service	-629.1	254.5	-2.472	0.01385

Tabela 1: Fitting linear model: salary ~ yrs.since.phd + yrs.service

Observations	Residual Std. Error	R^2	Adjusted R^2
397	27357	0.1883	0.1842

Jak widzimy, współczynniki regresji dla obu naszych predyktorów są istotnie różne od 0 (dla $\alpha = 0.05$). Widzimy również, że kontrolując liczbę lat, które upłynęły od uzyskania doktoratu, relacja między stażem a wynagrodzeniem jest negatywna. Każdy rok stażu zmniejsza (nieznacznie) wynagrodzenie!

Współczynnik R^2 dla naszego modelu wynosi 0.18, co jest wartością umiarkowanie wysoką jak na nauki społeczne. Sprawdźmy jeszcze, czy nasze predyktory nie są za bardzo współliniowe – mamy podstawy sądzić, że staż oraz lata, które upłynęły od uzyskania tytułu doktora mogą być ze sobą bardzo silnie skorelowane. Jak Państwo pamiętają, możemy w tym wypadku posłużyć się wartością VIF.

```
pander(vif(fit))
```

yrs.since.phd	yrs.service
5.796	5.796

Nasze zmienne objaśniające są do pewnego stopnia współliniowe (VIF w okolicach 6), ale nie jest to jakaś wielka tragedia (szczególnie w przykładzie mającym funkcje przede wszystkim pedagogiczne), więc nie przejmujemy się tym tutaj. Warto jednak mieć tę wartość w pamięci, bo do problemu współliniowości będziemy jeszcze wracać.

Regresja z interakcjami

Spróbujmy zastanowić się jakie mamy możliwości relacji między stażem, latami od uzyskania doktoratu jako zmiennymi objaśniającymi a wynagrodzeniem. W najprostszym wypadku efekt może być kumulatywny (suma efektów). Gdybyśmy mieli do czynienia z efektami o tym samym znaku, moglibyśmy zaobserwować synergistyczny efekt. W przypadku efektów o przeciwnym znaku moglibyśmy mówić o interakcji buforującej, w której efekty mają przeciwny znak i osłabiają się wzajemnie. Trzecim wzorcem interakcji jest interakcja antagonistyczna, w której oba predyktory wpływają na zmienną objaśnianą w ten sam sposób, ale interakcja ma przeciwny znak.

Zobaczmy, z którym z tych trzech ważnych wzorców mamy do czynienia w naszym przypadku. W R możemy zdecydować, czy chcemy badać również interakcje, używając zamiast znaku dodawania (+) znaku mnożenia (*).

```
fit <- lm(salary ~ yrs.since.phd * yrs.service, data = Salaries)
pander(summary(fit))
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	70155	3472	20.21	9.036e-63
yrs.since.phd	2194	246.9	8.889	2.26e-17
yrs.service	1692	356.3	4.75	2.852e-06
yrs.since.phd:yrs.service	-64.62	7.487	-8.63	1.537e-16

Tabela 2: Fitting linear model: salary ~ yrs.since.phd * yrs.service

Observations	Residual Std. Error	R^2	Adjusted R^2
397	25115	0.3177	0.3125

Zanim przejdziemy do interpretacji interakcji, to, żeby lepiej uzmysłowić sobie jak działają, spróbujmy ręcznie stworzyć zmienną, która będzie reprezentowała interakcje tych dwóch czynników. W kolumnie `interaction_term` w ramce danych umieścimy iloczyn wartości z kolumn `yrs.since.phd` oraz `yrs.service` a następnie stworzymy model (bez używania `*`), który będzie tę interakcję uwzględniał.

```
Salaries$interaction_term <- Salaries$yrs.since.phd * Salaries$yrs.service
fit <- lm(salary ~ yrs.since.phd + yrs.service + interaction_term, data =
  ↪ Salaries)
pander(summary(fit))
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	70155	3472	20.21	9.036e-63
yrs.since.phd	2194	246.9	8.889	2.26e-17
yrs.service	1692	356.3	4.75	2.852e-06
interaction_term	-64.62	7.487	-8.63	1.537e-16

Tabela 3: Fitting linear model: salary ~ yrs.since.phd + yrs.service + interaction_term

Observations	Residual Std. Error	R^2	Adjusted R^2
397	25115	0.3177	0.3125

Jak widzimy, uzyskaliśmy takie same wartości dla współczynników regresji, R^2 oraz taki sam wynik testów statystycznych. Pokazuje to, że możemy traktować interakcje jako wprowadzenie do modelu dodatkowej zmiennej, będącej iloczynem tych zmiennych, których interakcje chcemy zbadać. R ułatwia nam jednak sprawę i możemy po prostu napisać `*`.

Interpretacja interakcji

Wyświetlmy jeszcze raz podstawowe informacje dotyczącego naszego modelu z interakcją.

```
fit <- lm(salary ~ yrs.since.phd * yrs.service, data = Salaries)
pander(summary(fit))
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	70155	3472	20.21	9.036e-63
yrs.since.phd	2194	246.9	8.889	2.26e-17
yrs.service	1692	356.3	4.75	2.852e-06
yrs.since.phd:yrs.service	-64.62	7.487	-8.63	1.537e-16

Tabela 4: Fitting linear model: salary ~ yrs.since.phd * yrs.service

Observations	Residual Std. Error	R^2	Adjusted R^2
397	25115	0.3177	0.3125

Jak widzimy mamy w naszym modelu trzy zmienne: `yrs.since.phd`, `yrs.service` oraz interakcję czyli `yrs.since.phd:yrs.service`. Dla wszystkich z nich przekroczyliśmy próg statystycznej istotności. Jak je jednak zinterpretować? Wprowadzenie interakcji sprawia, że sytuacja się nieco komplikuje.

Współliniowość i centrowanie zmiennych

Jak pamiętamy, przy wielokrotnej regresji bez interakcji, nie wszystkie transformacje predyktorów wpływały na współczynniki regresji. Możemy to łatwo przetestować. Wycentrujemy sobie nasze predyktory, to znaczy dla każdej zmiennej od każdej wartości odejmiemy średnią dla danej zmiennej. Dzięki temu średnia dla każdego z predyktorów wynosić będzie 0.

```
Salaries$yrs.since.phd_c <- Salaries$yrs.since.phd -
  ↪ mean(Salaries$yrs.since.phd)
Salaries$yrs.service_c <- Salaries$yrs.service - mean(Salaries$yrs.service)
```

Teraz zobaczmy, czy wycentrowanie zmiennych wpływa na współczynniki regresji przy modelu bez interakcji.

```
fit_zwykly <- lm(salary ~ yrs.since.phd + yrs.service, data = Salaries)
fit_wycentrowany <- lm(salary ~ yrs.since.phd_c + yrs.service_c, data =
  ↪ Salaries)
pander(summary(fit_zwykly))
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	89912	2844	31.62	3.811e-110
yrs.since.phd	1563	256.8	6.086	2.754e-09
yrs.service	-629.1	254.5	-2.472	0.01385

Tabela 5: Fitting linear model: salary ~ yrs.since.phd + yrs.service

Observations	Residual Std. Error	R^2	Adjusted R^2
397	27357	0.1883	0.1842

```
pander(summary(fit_wycentrowany))
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	113706	1373	82.82	2.597e-251
yrs.since.phd_c	1563	256.8	6.086	2.754e-09
yrs.service_c	-629.1	254.5	-2.472	0.01385

Tabela 6: Fitting linear model: salary ~ yrs.since.phd_c + yrs.service_c

Observations	Residual Std. Error	R^2	Adjusted R^2
397	27357	0.1883	0.1842

Jak widzimy, zmienił się tylko wyraz wolny regresji, podczas gdy współczynniki kierunkowe dla wszystkich predyktorów pozostały te same. Nie powinno nas to dziwić. Co jednak, gdy zrobimy to samo dla modelu z interakcjami?

```
fit_zwykly <- lm(salary ~ yrs.since.phd * yrs.service, data = Salaries)
fit_wycentrowany <- lm(salary ~ yrs.since.phd_c * yrs.service_c, data =
  ↪ Salaries)
pander(summary(fit_zwykly))
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	70155	3472	20.21	9.036e-63
yrs.since.phd	2194	246.9	8.889	2.26e-17
yrs.service	1692	356.3	4.75	2.852e-06
yrs.since.phd:yrs.service	-64.62	7.487	-8.63	1.537e-16

Tabela 7: Fitting linear model: salary ~ yrs.since.phd * yrs.service

Observations	Residual Std. Error	R^2	Adjusted R^2
397	25115	0.3177	0.3125

```
pander(summary(fit_wycentrowany))
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	123533	1699	72.73	4.563e-230
yrs.since.phd_c	1056	243	4.346	1.764e-05
yrs.service_c	250.5	254.9	0.9829	0.3262
yrs.since.phd_c:yrs.service_c	-64.62	7.487	-8.63	1.537e-16

Tabela 8: Fitting linear model: salary ~ yrs.since.phd_c * yrs.service_c

Observations	Residual Std. Error	R^2	Adjusted R^2
397	25115	0.3177	0.3125

Widzimy, że stało się coś nieco nieoczekiwanego! Wartość R^2 pozostała taka sama, ale zmieniły się współczynniki regresji, co poskutkowało również zmianą tego, *które z nich są statystycznie istotne*. Dlaczego tak się stało? Żeby zrozumieć dlaczego wycentrowanie ma w tym wypadku wpływ na współczynniki regresji, musimy najpierw dobrze zrozumieć co one znaczą dla modeli z interakcjami.

Kiedy efekty poszczególnych predyktorów są kumulatywne (tzn. nie mamy modelu z interakcjami), to współczynniki regresji dla każdego z predyktorów jest taki sam niezależnie od wartości innych predyktorów (efekt pierwszego rzędu, *first order effect*). Sytuacja zmienia się jednak kiedy w naszym modelu uwzględniliśmy interakcje. Wówczas współczynniki regresji dla poszczególnych wchodzących w interakcje zmiennych moderowane są przez wartości tych zmiennych, z którymi dana zmienna wchodzi w interakcje.

Upraszczając trochę sprawę możemy powiedzieć, że skutkiem uwzględnienia w modelu interakcji jest to, że współczynnik regresji dla `yrs.since.phd` będzie różny w zależności od tego, czy `yrs.service` będzie na przykład wynosić 5 czy 15.

Do czego w takim razie odnoszą się wartości z kolumny `Estimate` dla naszych pojedynczych predyktorów? Są to wartości współczynników regresji dla efektów pierwszego rzędu kiedy wartości pozostałych predyktorów wynoszą 0. Dlatego właśnie zmieniły się kiedy wycentrowaliśmy nasze zmienne! 0 wszak oznaczało przed wycentrowaniem zero czegoś (lat pracy, lat stażu), teraz zaś oznacza średnią w naszej próbie (średnia lat pracy, średnia stażu). Mamy więc rozwiązanie zagadki!

Teraz możemy więc zinterpretować współczynniki regresji dla poszczególnych predyktorów. Możemy powiedzieć, że dla średniej wartości stażu wyrażonego w latach każdy rok który upłynął od uzyskania tytułu doktora zwiększa średnio wynagrodzenie o 1056 dolarów rocznie!

Dodatkowo możemy wreszcie powiedzieć, z którym wzorcem interakcji z omówionych wcześniej, mamy do czynienia. Oba nasze predyktory działają w tym samym kierunku, ale *interakcja ma przeciwny znak*. Zinterpretować możemy to tak, że oba rozważane przez nas czynniki (lata pracy i lata od doktoratu) mają pozytywny wpływ na wynagrodzenie, ale wpływ każdego z nich zmniejsza się wraz ze wzrostem drugiego! Dodatkowo warto zauważyć, że nasz model z interakcją ma sporo wyższy R^2 niż model bez interakcji.

Widzimy teraz dlaczego warto centrować nasze predyktory - możemy dzięki temu podać sensowną interpretację współczynników regresji dla efektów pierwszego rzędu. Gdybyśmy tego nie zrobili, to powiedzielibyśmy, że dla wartości stażu wyrażonego w latach również zero, każdy rok który upłynął od uzyskania tytułu doktora zwiększa średnio wynagrodzenie o 2194 dolary rocznie. To nie jest zbyt sensowna wartość zważywszy na to, że trudno mieć staż wynoszący 0 lat. Sytuacja wygląda jeszcze bardziej absurdalnie w przypadku takich zmiennych jak wiek, kiedy mówilibyśmy coś o sytuacji, w której ktoś ma 0 lat!

Dodatkowo w przypadku wycentrowania naszych predyktorów możemy zinterpretować współczynniki regresji dla efektów pierwszego rzędu w jeszcze jeden sposób. Reprezentują one wówczas *średni efekt danego predyktora* w całym zakresie danego predyktora, co jest raczej bliższe temu, jak interpretujemy te współczynniki w modelu bez interakcji. Ponadto, kiedy rozważamy średni efekt danego predyktora, łatwiej nam zinterpretować wynik testu statystycznego. Gdybyśmy nie wycentrowali predyktorów, to testy statystyczne odnosiłyby się do efektów pierwszego rzędu dla wartości pozostałych predyktorów (w rozważanej interakcji) równej 0.

Generalnie rzecz biorąc podręczniki zajmujące się regresją rekomendują w takich wypadkach centrowanie predyktorów za wyjątkiem sytuacji, kiedy wartość 0 ma jakąś interesującą interpretację w odniesieniu do tego, co nasza zmienna mierzy.

Centrowanie zmiennych ma jeszcze jedną zaletę - pozwala wyeliminować do pewnego stopnia współliniowość zmiennych z modelu. Porównajmy wartości VIF dla zmiennych wycentrowanych i niewycentrowanych.

```
pander(vif(fit_wycentrowany))
```

there are higher-order terms (interactions) in this model
consider setting type = 'predictor'; see ?vif

yrs.since.phd_c	yrs.service_c	yrs.since.phd_c:yrs.service_c
6.155	6.899	1.317

```
pander(vif(fit_zwykly))
```

there are higher-order terms (interactions) in this model
consider setting type = 'predictor'; see ?vif

yrs.since.phd	yrs.service	yrs.since.phd:yrs.service
6.354	13.48	13.01

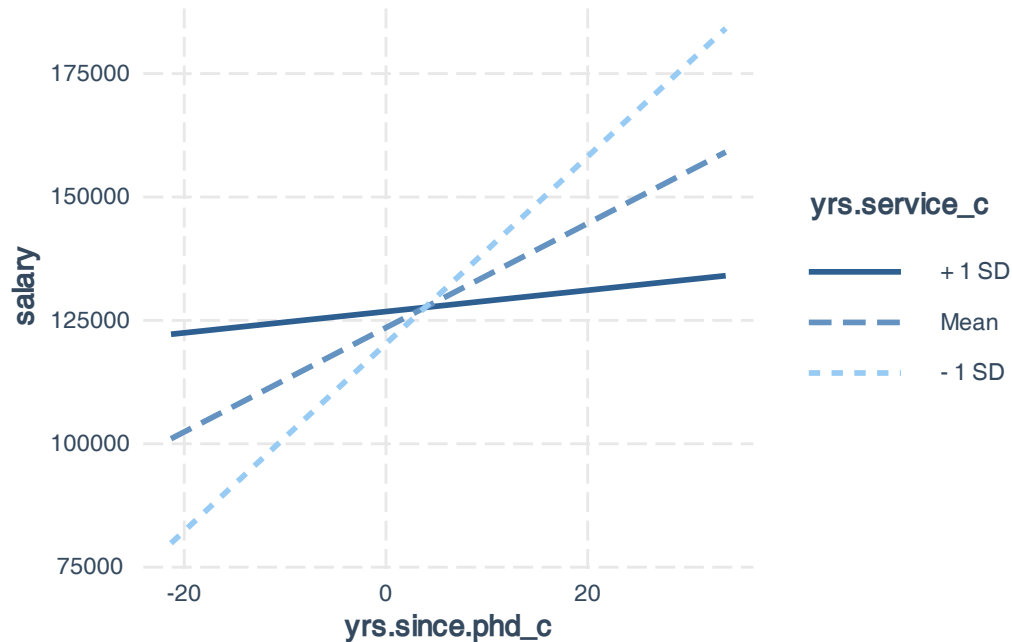
Widzimy, że centrowanie ograniczyło tutaj współliniowość między predyktorami a ich iloczynem, wynikającą z tego, gdzie na skalach znajduje się zero (tzw. nonessential collinearity). Dzięki temu wysokie wartości VIF związane z takim kodowaniem nie przesłaniają nam współliniowości między samymi predyktorami — tej centrowanie nie usuwa. Nie oznacza to poprawy dopasowania modelu ani dokładniejszego oszacowania interakcji.

Wizualizacja interakcji

Interakcje między zmiennymi można z powodzeniem zwizualizować. Dzięki przedstawieniu tej relacji na wykresie, łatwiej uświadomić sobie co tak naprawdę znaczą poszczególne liczby i jakie ma to przełożenie na relacje między zmiennymi. W pakiecie **interactions** znajduje się sporo narzędzi do analizy interakcji oraz ich wizualizacji. My skorzystamy z dostępnej w nim funkcji `interact_plot`.

Funkcja jako pierwszy argument przyjmuje model, który stworzyliśmy za pomocą funkcji `lm`, prześlemy jej również dwa dodatkowe argumenty - którą zmienną na użytek wykresu chcemy potraktować jako predyktor, a którą jako moderator. My chcemy dowiedzieć się, jak zmienia się relacja między wynagrodzeniami a czasem, który upłynął od uzyskania doktoratu wraz ze zmianą stażu zatrudnienia.

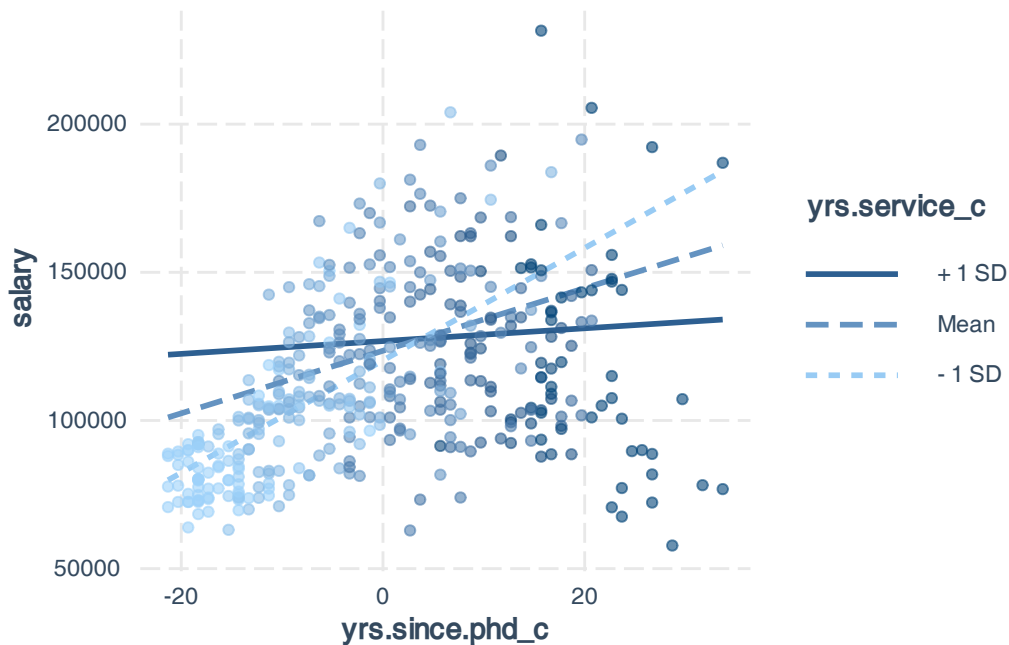
```
library(interactions)
interact_plot(fit_wycentrowany,
              pred = yrs.since.phd_c,
              modx = yrs.service_c)
```



Jak widzimy funkcja `interact_plot` wybrała domyślnie 3 poziomy naszego moderatora (`yrs.service_c`), dla których na wykres nałożona jest linia regresji. Są to: średnia, -1 odchylenie standardowe od średniej oraz +1 odchylenie standardowe od średniej. Widzimy teraz na czym polega moderowanie efektu pierwszego rzędu dla jednej zmiennej przez wartość drugiej zmiennej. Wraz ze wzrostem wartości naszego moderatora (stażu), linia regresji jest coraz bardziej pozioma.

Wykres tworzony za pomocą `interact_plot` możemy dostosowywać do swoich potrzeb. Poniżej znajduje się przykład, gdzie oprócz linii regresji dla trzech poziomów `yrs.service_c` nanieśliśmy na wykres również punkty (wykres rozrzutu).

```
interact_plot(fit_wycentrowany,
              pred = yrs.since.phd_c,
              modx = yrs.service_c,
              plot.points = TRUE, point.shape = TRUE)
```



Appendix - współczynniki regresji dla różnych poziomów moderatora

Na pewno zadali sobie Państwo pytanie, jak obliczyć wartość współczynnika regresji dla różnych wartości zmiennej moderującej w danej interakcji. Można to dość łatwo zrobić na piechotę (jak? to już zadanie do zastanowienia się dla Państwa!), ale funkcja `sim_slopes` z pakietu `interactions` zrobi to za nas (nawet umie sama centrować zmienne!).

W przypadku, gdy interesują nas wartości dla średniej, +1 odchylenia standardowego i -1 odchylenia standardowego używamy jej bardzo podobnie jak `interact_plot`.

```
sim_slopes(fit_zwykly,
  pred = yrs.since.phd,
  modx = yrs.service)
```

JOHNSON-NEYMAN INTERVAL

When `yrs.service` is OUTSIDE the interval [25.90, 44.67], the slope of `yrs.since.phd` is $p < .05$.

Note: The range of observed values of `yrs.service` is [0.00, 60.00]

SIMPLE SLOPES ANALYSIS

Slope of yrs.since.phd when yrs.service = 4.608586 (- 1 SD):

Est.	S.E.	t val.	p
1896.50	238.92	7.94	0.00

Slope of yrs.since.phd when yrs.service = 17.614610 (Mean):

Est.	S.E.	t val.	p
1056.09	242.98	4.35	0.00

Slope of yrs.since.phd when yrs.service = 30.620633 (+ 1 SD):

Est.	S.E.	t val.	p
215.68	282.76	0.76	0.45

Jeżeli interesują nas jakieś specyficzne wartości moderatora, to możemy je przekazać w argumentcie `modx.values`. Dla przykładu wybrałem wartości `yrs.service` równe 5, 30 i 55, a `sim_slopes` obliczyła współczynniki regresji dla `yrs.since.phd` oraz przeprowadziła dla nich test statystyczny ($H_0 : B = 0$, czyli czy dany współczynnik jest statystycznie różny od 0).

```
sim_slopes(fit_zwykly,
  pred = yrs.since.phd,
  modx = yrs.service,
  modx.values = c(5,30,55))
```

JOHNSON-NEYMAN INTERVAL

When yrs.service is OUTSIDE the interval [25.90, 44.67], the slope of yrs.since.phd is $p < .05$.

Note: The range of observed values of yrs.service is [0.00, 60.00]

SIMPLE SLOPES ANALYSIS

Slope of yrs.since.phd when yrs.service = 5.00:

Est.	S.E.	t val.	p
1871.20	238.46	7.85	0.00

Slope of yrs.since.phd when yrs.service = 30.00:

Est.	S.E.	t val.	p
255.78	280.23	0.91	0.36

Slope of yrs.since.phd when yrs.service = 55.00:

Est.	S.E.	t val.	p
-1359.64	412.63	-3.30	0.00