UNIVERSITY OF OSLO

NOR-EDEN

Introducing noreden

An R package for sustainable diet discovery

Chi Zhang

OCBE (Oslo Center for Biostatistics and Epidemiology) & Dept of Nutrition, University of Oslo

2023-11-24 NorEden Annual Meeting

Background

Joined 2023.04 as 50% researcher

Statistician and R programmer

Work Package II

Collaborators: Julie, Lene, Arnoldo (Dept of Biostatistics)

This talk: introduce a new R package **noreden** (work-in-progress)

Facilitate sustainable diet discovery, when we have information of the existing diet such as nutrition, environmental impact, etc

What is an R package?

What is an R package?

Open source software

What is an R package?

Open source software

Free and open to everyone to download and use

What is an R package?

Open source software

Free and open to everyone to download and use

Source code is shared in **public** repository

What is an R package?

Open source software

Free and open to everyone to download and use

Source code is shared in **public** repository

Open to **feedback and collaboration** by others

What is an R package?

Open source software

Free and open to everyone to download and use

Source code is shared in **public** repository

Open to **feedback and collaboration** by others

Development and release (new ‘versions’) can keep going after the paper is published

What is an R package?

Open source software

Free and open to everyone to download and use

Source code is shared in **public** repository

Open to **feedback and collaboration** by others

Development and release (new ‘versions’) can keep going after the paper is published

What is an R package?

Open source software

Free and open to everyone to download and use

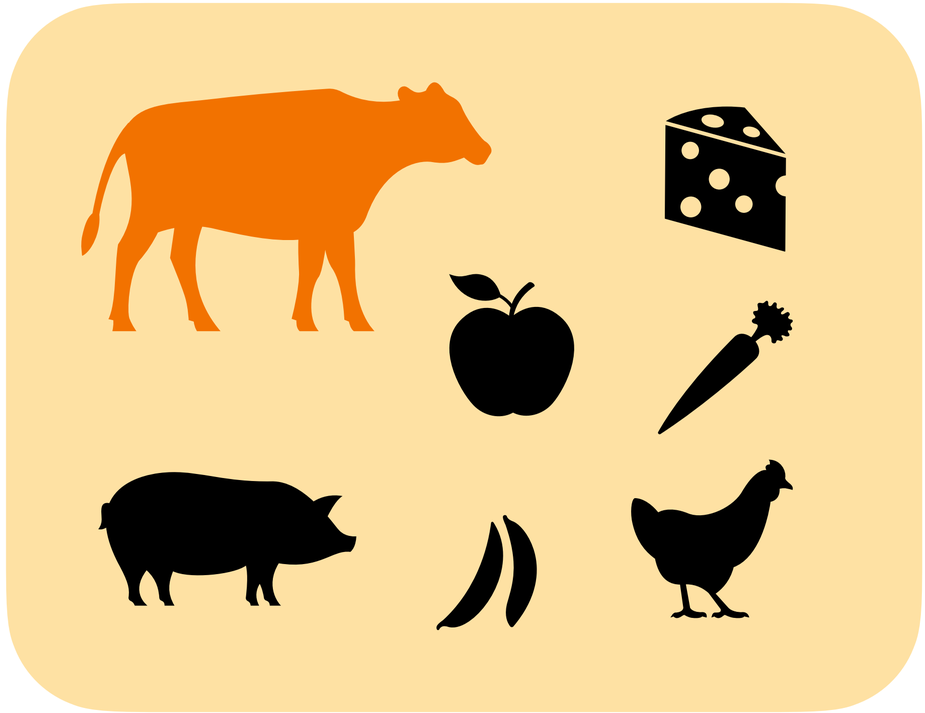
Source code is shared in **public** repository

Open to **feedback and collaboration** by others

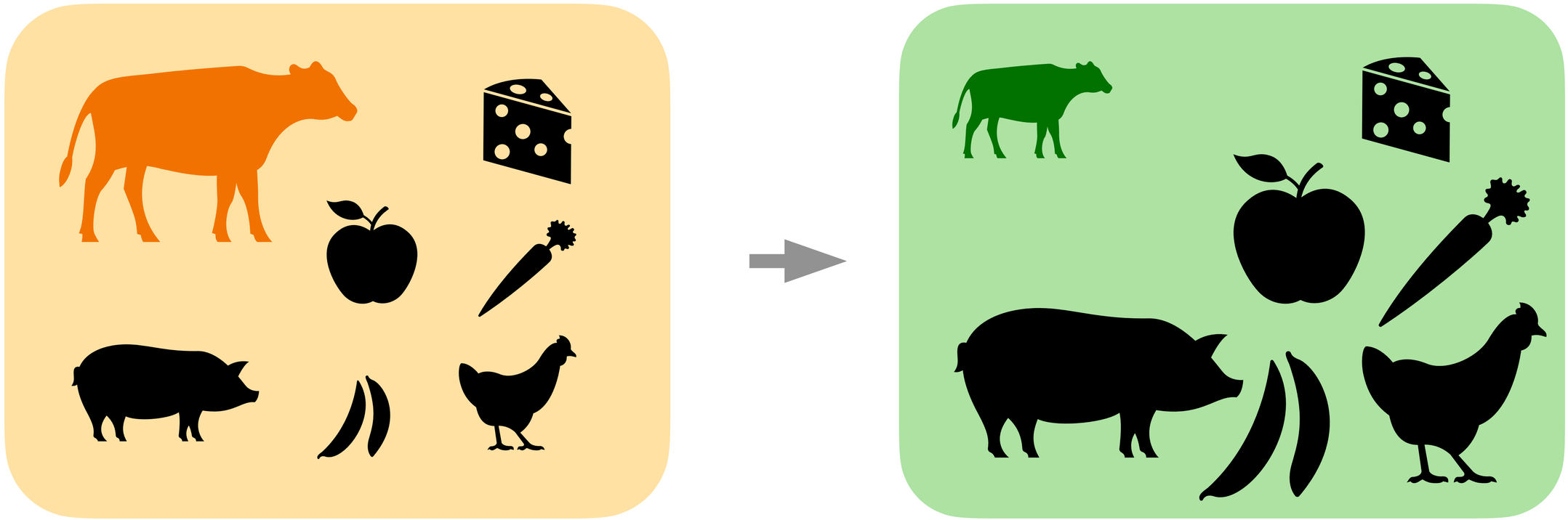
Development and release (new ‘versions’) can keep going after the paper is published

Publishing source code and implementation details is getting more and more important!

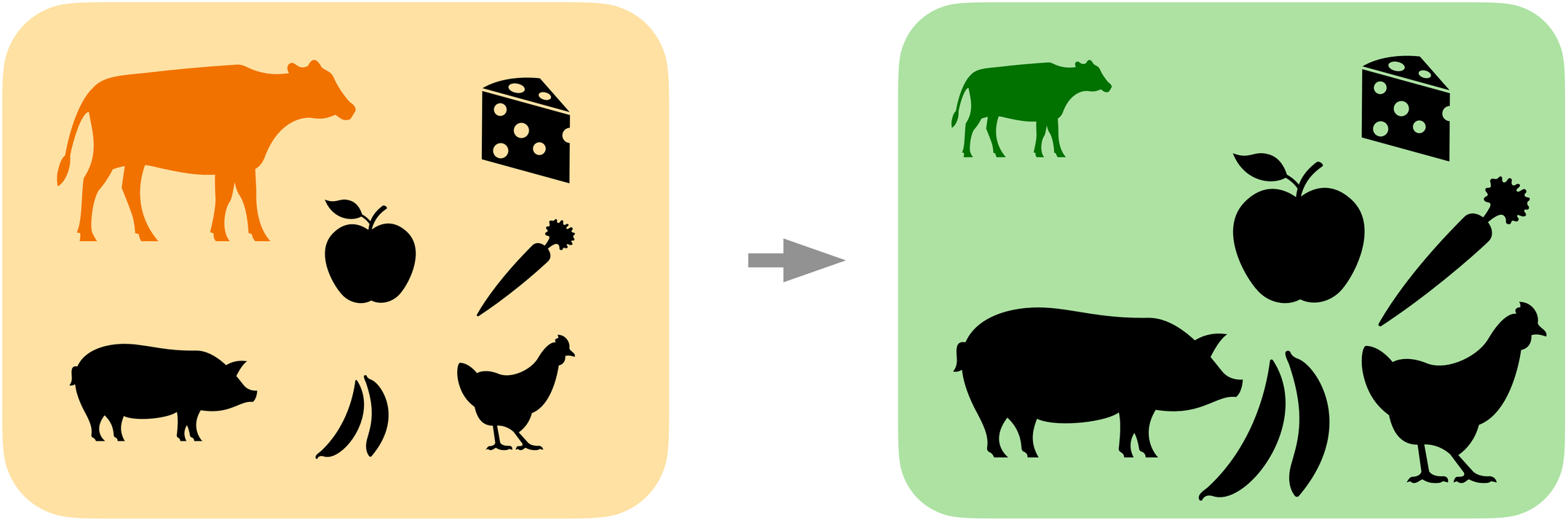
Motivation



Motivation



Motivation



Sustainable transition:
Healthy, more environmental friendly, yet acceptable by most people

Optimization setup

Healthy

More environmental friendly

Acceptable by most people

(Current diet is based on dietary survey from 1700+ subjects)

Optimization setup

Healthy	falls within a healthy range for nutrition outcomes
More environmental friendly	
Acceptable by most people	

(Current diet is based on dietary survey from 1700+ subjects)

Optimization setup

Healthy	falls within a healthy range for nutrition outcomes
More environmental friendly	produces less green house gas (or other outcomes)
Acceptable by most people	

(Current diet is based on dietary survey from 1700+ subjects)

Optimization setup

Healthy	falls within a healthy range for nutrition outcomes
More environmental friendly	produces less green house gas (or other outcomes)
Acceptable by most people	is not too far from the current one, and make practical sense

(Current diet is based on dietary survey from 1700+ subjects)

Optimization setup

Healthy	falls within a healthy range for nutrition outcomes
More environmental friendly	produces less green house gas (or other outcomes)
Acceptable by most people	is not too far from the current one, and make practical sense

	food	intake	energy	protein	fat	carbs	sugar	alcohol	ghge
1:	Bread	175.4	10.696	0.091	0.030	0.441	0.002	0	0.001
2:	Vegetables	154.6	1.565	0.015	0.008	0.050	0.005	0	0.001
3:	Red meat	117.6	8.342	0.173	0.139	0.014	0.000	0	0.013

current diet (gram)

produce (per gram)

(Current diet is based on dietary survey from 1700+ subjects)

Optimization setup

Healthy	falls within a healthy range for nutrition outcomes
More environmental friendly	produces less green house gas (or other outcomes)
Acceptable by most people	is not too far from the current one, and make practical sense

	food	intake	energy	protein	fat	carbs	sugar	alcohol	ghge
1:	Bread	175.4	10.696	0.091	0.030	0.441	0.002	0	0.001
2:	Vegetables	154.6	1.565	0.015	0.008	0.050	0.005	0	0.001
3:	Red meat	117.6	8.342	0.173	0.139	0.014	0.000	0	0.013

current diet (gram)

produce (per gram)

3 food example:

175g, 154g, 117g -> ?

Reduce ghge by 10%?

(Current diet is based on dietary survey from 1700+ subjects)

Optimization setup

Healthy	falls within a healthy range for nutrition outcomes
More environmental friendly	produces less green house gas (or other outcomes)
Acceptable by most people	is not too far from the current one, and make practical sense

	food	intake	energy	protein	fat	carbs	sugar	alcohol	ghge
1:	Bread	175.4	10.696	0.091	0.030	0.441	0.002	0	0.001
2:	Vegetables	154.6	1.565	0.015	0.008	0.050	0.005	0	0.001
3:	Red meat	117.6	8.342	0.173	0.139	0.014	0.000	0	0.013

current diet (gram)

produce (per gram)

3 food example:

175g, 154g, 117g -> ?

Reduce ghge by 10%?

(Current diet is based on dietary survey from 1700+ subjects)

noreden package

noreden 0.1.0 Get started Reference Search for

noreden

The package `noreden` provides user-friendly functions and tools for nutrition researcher to explore diet design, under various nutritional and environmental constraints.

Scope

- New diet discovery based on current Norwegian dietary survey, subject to nutritional and environmental impact inequality constraints
- Quadratic Programming (QP) implemented by `nloptr`

Main Features

- Flexible constraint specification via coefficient
- Tidy results with tabular and visual presentation

Installation

Development

The package is being actively developed. If you wish to test and contribute to the project, you can download from here,

```
devtools::install_github("andreaczhang/noreden")
```

Get Started

Read the introductory vignette for more information.

Links

Browse source code

Report a bug

License

Full license

MIT + file LICENSE

Citation

Citing noreden

Developers

Chi Zhang
Author, maintainer

Julie Marie Lengle
Author

Dev status

lifecycle experimental

Structure:

Process input

Set constraints

Find new diet

Present results

Implemented via “functions”

noreden package

Step 1: process input

12 foods
5 outputs (energy, protein, carbs, fat, ghge)

Need to know what the current intakes are; and how much they “contribute”; individ + total

noreden package

Step 1: process input

12 foods
5 outputs (energy, protein, carbs, fat, ghge)

Need to know what the current intakes are; and how much they “contribute”; individ + total

```
diet_selected <- select_diet(
  data_diet = all_diet,
  tag_food = tag_food_12)

diet_selected
#> # A tibble: 12 × 4
#>    food_name              intake_mean intake_lwr intake_upr
#>    <chr>                        <dbl>      <dbl>      <dbl>
#>  1 Bread                        188.       18.8       344.
#>  2 Other grains                  48.3       4.83      156.
#>  3 Potatoes                      72.8       7.28      231.
#>  4 Vegetables                   166.       16.6       420.
#>  5 Fruit, berries               184.       18.4       553.
#>  6 Red meat                     126.       12.6       300.
#>  7 Fish                          74.6       7.46      303.
#>  8 Eggs                          26.4       2.64      112.
#>  9 Milk, yoghurt                329.       32.9       901.
#> 10 Cheese                       46.6       4.66      122.
#> 11 Butter, margarine, oil       32.3       3.23       71.4
#> 12 Sugar, sweets                18.1       1.81       66
```

noreden package

Step 1: process input

12 foods
5 outputs (energy, protein, carbs, fat, ghge)

Need to know what the current intakes are; and how much they “contribute”; individ + total

```
diet_selected <- select_diet(
  data_diet = all_diet,
  tag_food = tag_food_12)

diet_selected
#> # A tibble: 12 × 4
#>    food_name              intake_mean intake_lwr intake_upr
#>    <chr>                        <dbl>      <dbl>      <dbl>
#>  1 Bread                        188.       18.8       344.
#>  2 Other grains                  48.3       4.83      156.
#>  3 Potatoes                      72.8       7.28      231.
#>  4 Vegetables                   166.       16.6       420.
#>  5 Fruit, berries               184.       18.4       553.
#>  6 Red meat                     126.       12.6       300.
#>  7 Fish                          74.6       7.46      303.
#>  8 Eggs                          26.4       2.64      112.
#>  9 Milk, yoghurt                329.       32.9       901.
#> 10 Cheese                       46.6       4.66      122.
#> 11 Butter, margarine, oil       32.3       3.23       71.4
#> 12 Sugar, sweets                18.1       1.81       66
```

```
cpu_selected <- select_perunit(
  data_perunit_contrib = contrib_per_unit,
  tag_food = tag_food_12,
  tag_outcome = tag_outcome_5)
cpu_selected
#> # A tibble: 12 × 6
#>    food_name              energy protein   carbs     fat     ghge
#>    <chr>                   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>
#>  1 Bread                   10.7  0.0912  0.441   0.0302  0.00107
#>  2 Other grains            14.0  0.1     0.607   0.0422  0.00235
#>  3 Potatoes                 3.79 0.0206  0.178   0.00737 0.00037
#>  4 Vegetables               1.57 0.0149  0.0498  0.00841 0.00103
#>  5 Fruit, berries           2.73 0.00758 0.134   0.00408 0.00072
#>  6 Red meat                 8.34 0.173   0.0136  0.139   0.0129
#>  7 Fish                     6.09 0.170   0.0245  0.0748  0.00311
#>  8 Eggs                     6.18 0.130   0.00407 0.106   0.00215
#>  9 Milk, yoghurt            1.98 0.0359  0.0559  0.0111  0.00143
#> 10 Cheese                  13.5  0.217   0.0484  0.242   0.0103
#> 11 Butter, margarine, oil  23.4  0.0133  0.0233  0.615   0.00467
#> 12 Sugar, sweets           18.0  0.0533  0.609   0.178   0.00387
```

noreden package

Step 2: set constraints

Reduce ghge to its 90% previous level

```
tc <- compute_total_contrib(
  data_diet = diet_selected,
  data_perunit_contrib = cpu_selected)

tc
#> $total_contrib
#>   tag_outcome total_contrib
#> 1      energy   7762.508053
#> 2     protein     88.576337
#> 3       carbs    194.438480
#> 4         fat     74.404123
#> 5        ghge      3.739236
#>
#> $tag_food
#>  [1] "Bread"                "Other grains"           "Potatoes"
#>  [4] "Vegetables"           "Fruit, berries"         "Red meat"
#>  [7] "Fish"                 "Eggs"                   "Milk, yoghurt"
#> [10] "Cheese"               "Butter, margarine, oil" "Sugar, sweets"
#>
#> $tag_outcome
#> [1] "energy"  "protein" "carbs"   "fat"     "ghge"
```

noreden package

Step 2: set constraints

Reduce ghge to its 90% previous level

```
tc <- compute_total_contrib(
  data_diet = diet_selected,
  data_perunit_contrib = cpu_selected)

tc
#> $total_contrib
#>   tag_outcome total_contrib
#> 1      energy   7762.508053
#> 2     protein     88.576337
#> 3       carbs    194.438480
#> 4         fat     74.404123
#> 5        ghge      3.739236
#>
#> $tag_food
#>  [1] "Bread"                "Other grains"           "Potatoes"
#>  [4] "Vegetables"           "Fruit, berries"         "Red meat"
#>  [7] "Fish"                 "Eggs"                   "Milk, yoghurt"
#> [10] "Cheese"               "Butter, margarine, oil" "Sugar, sweets"
#>
#> $tag_outcome
#> [1] "energy"  "protein" "carbs"   "fat"     "ghge"
```

```
# reduce ghge to 0.9
constr_coef_df_red <- reduce_constr(
  data_constr_coef = constr_coef_df,
  tag_outcome_reduce = 'ghge',
  coef_reduce = 0.9)
constr_coef_df_red
#>   tag_outcome coef_constrlwr coef_construpr
#> 1      energy           0.90            1.0
#> 2     protein           0.90            1.0
#> 3       carbs           0.90            1.0
#> 4         fat           0.90            1.0
#> 5        ghge           0.81            0.9
```

noreden package

Step 2: set constraints

Reduce ghge to its 90% previous level

```
tc <- compute_total_contrib(
  data_diet = diet_selected,
  data_perunit_contrib = cpu_selected)

tc
#> $total_contrib
#>   tag_outcome total_contrib
#> 1      energy   7762.508053
#> 2     protein     88.576337
#> 3       carbs    194.438480
#> 4         fat     74.404123
#> 5        ghge      3.739236
#>
#> $tag_food
#>  [1] "Bread"                "Other grains"          "Potatoes"
#>  [4] "Vegetables"           "Fruit, berries"        "Red meat"
#>  [7] "Fish"                 "Eggs"                  "Milk, yoghurt"
#> [10] "Cheese"               "Butter, margarine, oil" "Sugar, sweets"
#>
#> $tag_outcome
#> [1] "energy"  "protein" "carbs"   "fat"     "ghge"
```

```
# reduce ghge to 0.9
constr_coef_df_red <- reduce_constr(
  data_constr_coef = constr_coef_df,
  tag_outcome_reduce = 'ghge',
  coef_reduce = 0.9)
constr_coef_df_red
#>   tag_outcome coef_constrlwr coef_construpr
#> 1      energy           0.90            1.0
#> 2     protein           0.90            1.0
#> 3       carbs           0.90            1.0
#> 4         fat           0.90            1.0
#> 5        ghge           0.81            0.9
```

```
constr_val_reduce <- compute_constr(
  data_total_contrib = tc$total_contrib,
  data_constr_coef = constr_coef_df_red)
constr_val_reduce
#>   tag_outcome total_contrib coef_constrlwr coef_construpr  constr_lwr
#> 1      energy   7762.508053           0.90            1.0 6986.257248
#> 2     protein     88.576337           0.90            1.0   79.718703
#> 3       carbs    194.438480           0.90            1.0  174.994632
#> 4         fat     74.404123           0.90            1.0   66.963711
#> 5        ghge      3.739236           0.81            0.9    3.028781
#>    constr_upr
#> 1 7762.508053
#> 2   88.576337
#> 3  194.438480
#> 4   74.404123
#> 5    3.365312
```

noreden package

Step 3: find new diet

Inputs

current diet intake
outcome names
constraint values computed from before

Algorithm:
quadratic programming via nplotr

Minimize (diet_new, diet_current)
subject to

total energy <= upper bound of energy
total energy >= lower bound of energy
…

noreden package

Step 3: find new diet

Inputs

current diet intake
outcome names
constraint values computed from before

Algorithm:
quadratic programming via nplotr

Minimize (diet_new, diet_current)
subject to

total energy <= upper bound of energy
total energy >= lower bound of energy
…

```
res <- find_new_diet(diet0 = diet_selected$intake_mean,
                     diet0_upr = diet_selected$intake_upr,
                     diet0_lwr = diet_selected$intake_lwr,
                     tag_outcomes = tag_outcome_5,
                     constraint_val = constval$val,
                     print_runtime = T)
# collect result
new_diet <- return_new_diet(
  result_obj = res$run_optim,
  data_current_diet = diet_selected)
new_diet
#>                  food_name       new   current
#> 1                    Bread 187.11762 188.31866
#> 2             Other grains  45.56513  48.31437
#> 3                 Potatoes  72.37999  72.79364
#> 4               Vegetables 164.71925 165.98669
#> 5           Fruit, berries 183.25414 184.13142
#> 6                 Red meat 110.39761 126.26154
#> 7                     Fish  71.10191  74.61885
#> 8                     Eggs  24.33138  26.41185
#> 9            Milk, yoghurt 326.87133 328.64505
#> 10                 Cheese  35.17652  46.59652
#> 11 Butter, margarine, oil  30.50184  32.31694
#> 12          Sugar, sweets  14.33111  18.14473
```

noreden package

Step 4: present results

Validate whether new diet satisfies constraints

Compare change (gram, percent)

noreden package

Step 4: present results

Validate whether new diet satisfies constraints

Compare change (gram, percent)

```
# validate constraints
new_diet_validate <- validate_diet_contrib(data_new_diet = new_diet,
                      data_unit_contrib = cpu_selected,
                      data_constr = constr_val_reduce)

new_diet_validate
#>   tag_outcome total_contrib_new total_contrib coef_constrlwr coef_construpr
#> 1      energy           7269.81       7762.51           0.90            1.0
#> 2     protein             81.79         88.58           0.90            1.0
#> 3       carbs            188.66        194.44           0.90            1.0
#> 4         fat             66.96         74.40           0.90            1.0
#> 5        ghge              3.37          3.74           0.81            0.9
#>   constr_lwr constr_upr check deviation
#> 1    6986.26    7762.51    ok         0
#> 2      79.72      88.58    ok         0
#> 3     174.99     194.44    ok         0
#> 4      66.96      74.40    ok         0
#> 5       3.03       3.37    ok         0
```

noreden package

Step 4: present results

Validate whether new diet satisfies constraints

Compare change (gram, percent)

```
# validate constraints
new_diet_validate <- validate_diet_contrib(data_new_diet = new_diet,
                      data_unit_contrib = cpu_selected,
                      data_constr = constr_val_reduce)

new_diet_validate
#>   tag_outcome total_contrib_new total_contrib coef_constrlwr coef_construpr
#> 1      energy           7269.81       7762.51           0.90            1.0
#> 2     protein             81.79         88.58           0.90            1.0
#> 3       carbs            188.66        194.44           0.90            1.0
#> 4         fat             66.96         74.40           0.90            1.0
#> 5        ghge              3.37          3.74           0.81            0.9
#>   constr_lwr constr_upr check deviation
#> 1    6986.26    7762.51    ok         0
#> 2      79.72      88.58    ok         0
#> 3     174.99     194.44    ok         0
#> 4      66.96      74.40    ok         0
#> 5       3.03       3.37    ok         0
```

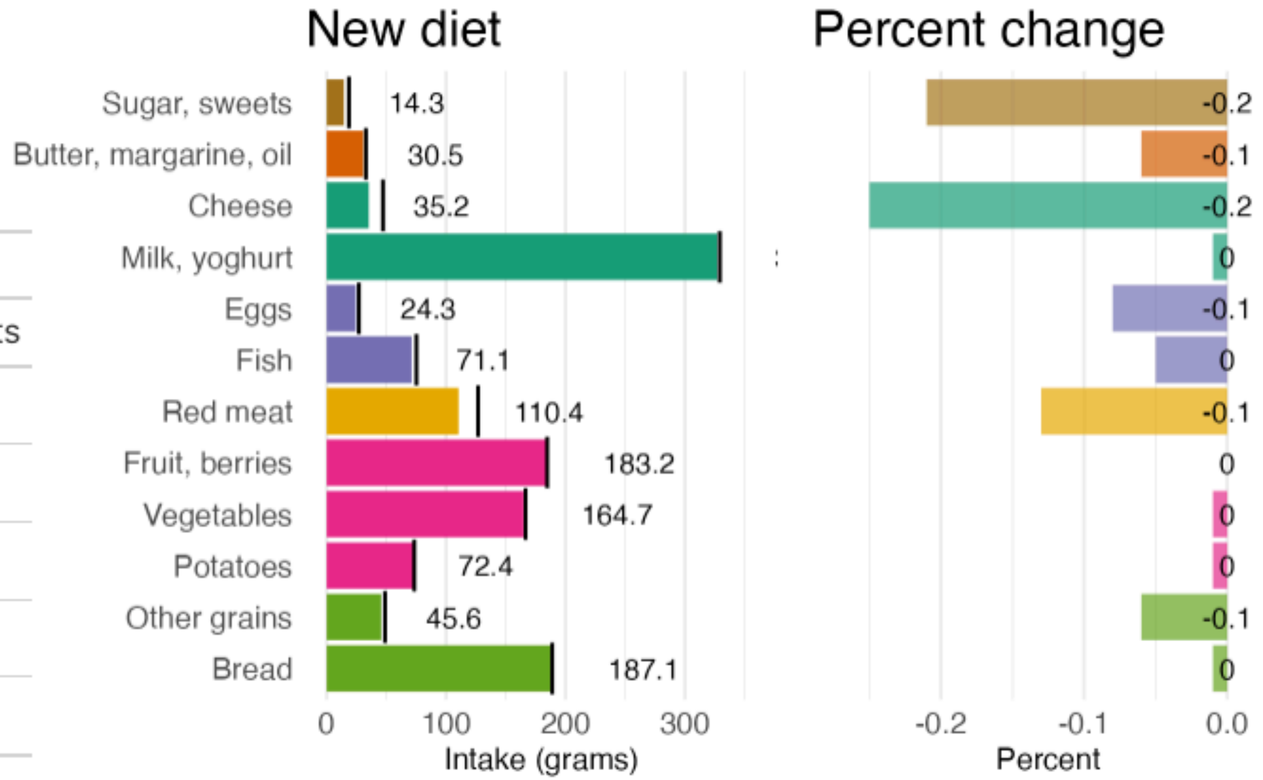
```
# compute difference
new_old_compare <- compare_new_diet(data_new_diet = new_diet,
                    data_current_diet = diet_selected)

new_old_compare
#>                  food_name    new current current_lwr current_upr abs_change
#> 1                    Bread 187.12  188.32       18.83       343.8      -1.20
#> 2              Other grains  45.57   48.31        4.83       155.7      -2.75
#> 3                  Potatoes  72.38   72.79        7.28       230.7      -0.41
#> 4                Vegetables 164.72  165.99       16.60       419.7      -1.27
#> 5           Fruit, berries 183.25  184.13       18.41       552.7      -0.88
#> 6                 Red meat 110.40  126.26       12.63       299.6     -15.86
#> 7                     Fish  71.10   74.62        7.46       302.9      -3.52
#> 8                     Eggs  24.33   26.41        2.64       111.6      -2.08
#> 9            Milk, yoghurt 326.87  328.65       32.86       900.6      -1.77
#> 10                  Cheese  35.18   46.60        4.66       121.6     -11.42
#> 11 Butter, margarine, oil  30.50   32.32        3.23        71.4      -1.82
#> 12           Sugar, sweets  14.33   18.14        1.81        66.0      -3.81
#>    perc_change
#> 1        -0.01
#> 2        -0.06
#> 3        -0.01
#> 4        -0.01
#> 5         0.00
#> 6        -0.13
#> 7        -0.05
#> 8        -0.08
#> 9        -0.01
#> 10       -0.25
#> 11       -0.06
#> 12       -0.21
```

noreden package

Step 4: present results

	Total contribution		Constraint	
Outcome	New diet	Current diet	Range	Comments
energy	**7267.99**	7762.51	6986.26-7762.51	Ok
protein	**81.75**	88.58	79.72-88.58	Ok
carbs	**188.59**	194.44	174.99-194.44	Ok
fat	**66.96**	74.40	66.96-74.40	Ok
ghge	**3.37**	3.74	3.03-3.37	Ok



Next steps

Improve documentation

Robust testing, fix existing bugs

Submit to CRAN (public repository), release version 1.0.0

Add more features (e.g. customizable input, optimization algorithm)