**Optimization of Dietary Nutrient Supplementation for Rational Rebalancing of Human Gut Microbiome**

Zihan Wang, Ray Huang and Sang Yoon Byun

**Definitions**

Amplicon Sequences Variants (ASVs): A cluster of closely related microbial species.  
Nutrient Impact (NI): A numerical value [0, 1] representing the anticipated propensity for a given ASV to benefit from supplementation of a specific nutrient.   
Nutrient Impact Matrix (NIM): An 𝑚×𝑛 matrix **A,** where 𝑚 is the number of ASVs, and 𝑛 is the number of nutrients, that aggregates the computed NI values for *n* nutrients over the entire set of ASVs. Each of **A**’s elements belong to [0, 1].  
Test Microbiome Sample (TMS): A *m*×1 vector **u,** representing the percentages of the total population number of the ASVs of the deviant sample.  
Normal Abundance Lower Bound: A vector **v**lowdenoting the lower bounds of normal state for every ASV.  
Normal Abundance Upper Bound: A vector **v**high denoting the upper bound of normal state for every ASV.  
Intervention:An *n*×1 vector **r**, with each value a positive real number, representing the nutritional intervention that increases nutrient abundance. The values in the vector represent the units of nutrients we provide in the dietary supplementation.   
Impact: A vector **b** = **A**×**r** representing how the intervention alters the given TMS.   
Failure Counts: An integer representing the number of ASVs that do not fall into the normal abundance range. Let *c* denote failure counts, where 𝑐 = 𝑚 − |{𝑖 | **v**𝑖*low*≤ (**u**𝑖+ **b**𝑖) / (+ 1≤𝑗≤𝑚 **b**j) ≤ **v**𝑖*high* , 1 ≤ 𝑖 ≤ 𝑚}|**The Approach**We consider our TMS as percentages of ASVs, but also associated with a value that tells the total amount of ASVs. As a preliminary step, we’ll compute the normal abundance range that each ASV in the TMS should fall into (see Preliminary Normal State Problem). We can then focus on finding the optimal subset of nutrients that makes the maximal number of ASVs fall into the normal states. We allow each selected nutrient to have multiplicities, as providing more nutrients can activate more ASVs. The given NIM table will decide how each ASV will be impacted by the nutrients. We consider the changes in ASV values as the dot product of the NIs of the ASV with the nutritional intervention.

**Dietary Nutrient Supplementation Optimization Problem:**

*Find a subset of nutrients that collectively minimizes the failure counts.*

**Input:** The NIM matrix **A**, a TMS **u**, and a desired number of nutrients used t.

**Output:** A vector **r,** such that **||r||0** t(where ||**r**||0 is the number of non-zero elements in **r**) that minimizes failure counts *c*.

**Preliminary Normal State Problem***Find a list of normal ASV abundance ranges for a particular TMS based on the RSC.*

Reference Sample Collection (RSC): A collection of vectors describing microbiome samples that are presumed to be normal.

**Input:** A Reference Sample Collection *RSC* and a test microbiome sample *TMS* **Output:** Two vectors **v**lowand **v**high of a given *TMS* that defines the region of normal ASV percentages.  
Note: We deliberately omitted the definition of a normal state here, hence this preliminary problem is not well-defined. This is because we want to consider this problem already solved and use **v**lowand **v**high in our main problem directly. There could be many definitions of a **v**lowand **v**high, and one possible heuristic we might use is for each ASV, we calculate the 10%, 90% quantile among the records as the low and high values.

**Timeline**

Week 4/5: Update problem formulation, work based on feedback received, and begin on implementation.

Week 6: Submit an extended description of the project and continue working on implementation.   
Week 7: Provide a summary of progress and preliminary results, and continue working on implementation.

Week 9: Deadline for 5-page long paper.