## Reshaping CGM data into Ax = y

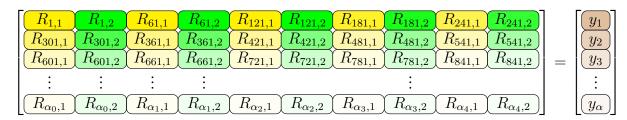
Assume the raw and glucose readings data has been trimmed and aligned already such that  $\alpha$  and  $\beta$  are known

Let  $R \in \mathbb{R}^{\beta \times n}$  denote the raw data<sup>1</sup> matrix where  $R_{i,j}$  is the measure of the  $j^{\text{th}}$  variable at the  $i^{\text{th}}$  second for  $i \leq \beta$  and  $j \leq n$ .

$$R = \begin{bmatrix} R_{1,1} & R_{1,2} & R_{1,3} & R_{1,4} & \dots & R_{1,n} \\ R_{2,1} & R_{2,2} & R_{2,3} & R_{2,4} & \dots & R_{2,n} \\ R_{3,1} & R_{3,2} & R_{3,3} & R_{3,4} & \dots & R_{3,n} \\ \vdots & \vdots & \vdots & & \vdots \\ R_{\beta,1} & R_{\beta,2} & R_{\beta,3} & R_{\beta,4} & \dots & R_{\beta,n} \end{bmatrix}$$

Now let  $y \in \mathbb{R}^{\alpha}$  be the vector of glucose measurements<sup>2</sup> which are recorded every 5 minutes (300 seconds) where  $y_k$  is the measure at the  $k^{\text{th}}$  5-minute interval for  $k \le \alpha << \beta$ . To have a system of equations of the form  $A_i x = y_i$ , the a row  $A_i$  must contain 5 minutes worth of rows from R. If we use all 300 possible rows (1 second per row) and all n variables from R then Ax = y will be over-determined since  $\alpha << 300n$ . Instead, we use  $\ell$  features and take a sample of  $\tau$  rows of R per minute such that  $5\ell\tau < \alpha$  and the system is under-determined.

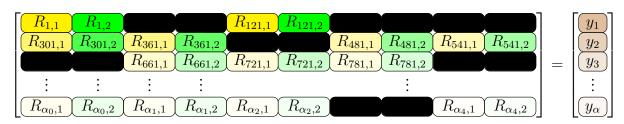
For example, assume  $\tau=1$  and we choose variables  $R_1$  and  $R_2$  ( $\ell=2$ ) to predict glucose measures. Then rows only have length 10 and Ax=y takes the form



where  $\alpha_i = 300\alpha + 61i$ .

## **Noise Masking**

Let  $N \in \mathbb{R}^{\beta}$  be the noise vector such that  $N_i$  is the recorded noise level at the  $i^{\text{th}}$  second and let  $\theta$  be the noise level we consider 'too noisy'. If  $N_i > \theta$  we wish to mask the tuple corresponding to the  $i^{\text{th}}$  second. The matrices returned from this masking algorithm will have the form



 $<sup>{}^{1}</sup>R$  is the concatenation of the "Breathing", "ECG", and "Summary" excel files

 $<sup>^{2}</sup>y$  is the "glucose" excel file