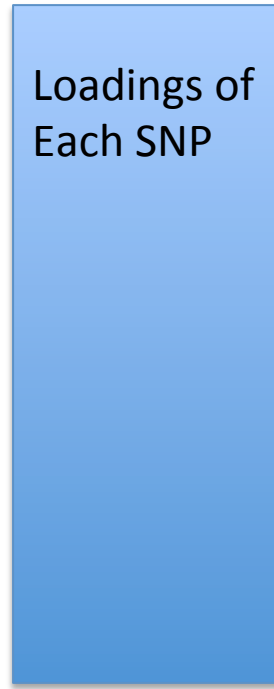




$N \times D$



$N \times K$

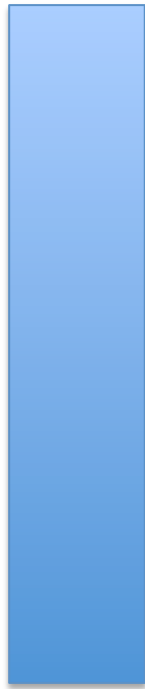


$K \times D$

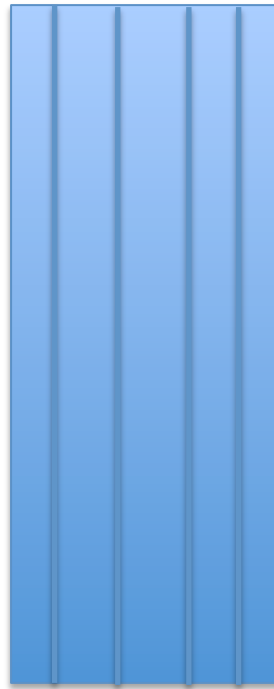
Factors:
Each Tissue Can be
active in more than
one Config

SFA

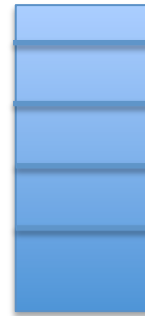
- Each SNP can be maximally loaded on 1 eigenconfig
- This means that it will be active only in tissues that are active in this eigenconfig
- Factor Represents the 'Eigenconfig' - i.e., the D Dimensional vector of directions active in a particular eigenconfig
- E.g., Factor 2 might consist of expression in the direction of tissue 1, Factor 4 might consist of expression in the direction of tissues 1 and 2, etc.
- We have hierarchically reduced K by estimating the proportional membership of each SNP in a config type



$X_{n \times d}$



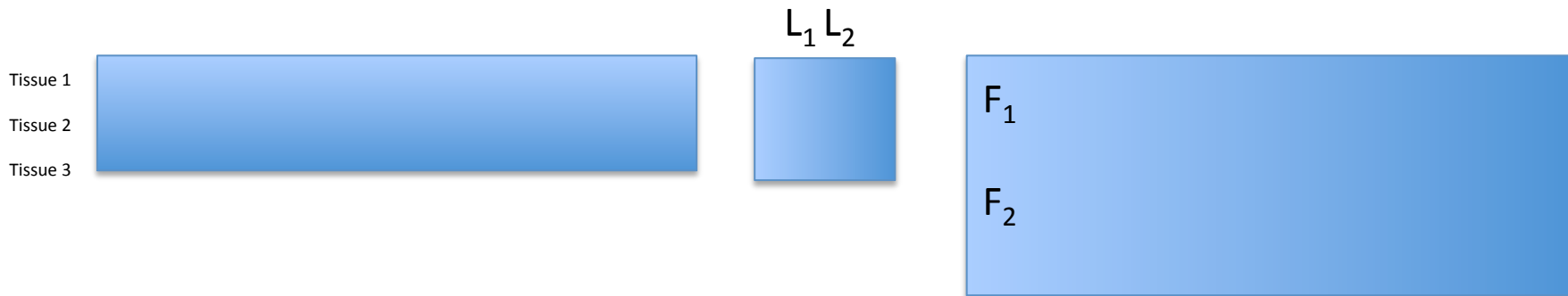
$U_{n \times k}$



V^t

- PCA
- Columns of U are linear combinations of tissues at a given SNP $U_{\cdot 1}$ represents direction of expression across the genome for eigentissue1
- Previously 'avg' given gene expression across tissues
- Now only the expression of a given gene in eigentissue1
- If we put sparse prior on this, it means, that a SNP can be active in only one 'eigenconfig'
- Rows of V^t are linear combinations of rows of X , each element of $V^{t[1,]}$ represents the genome-wide expression of eigentissue 1 in tissue 1
- Will only receive weight from SNPs active in eigentissue 1. Previously, 'average expression of all SNPs in tissue 1'
- Now expression of SNPs active in eigentissue1

Imagine $D = 3$ Tissues, $P = 10$ Gene-SNP Pair, 2 Factors



- #set possible loadings to be "sparse" (loaded on one factor each) means that each tissue can be loaded on one eigentissue each, which doesn't make sense because our aim is not to reduce the number of tissues but the number of ways in which a gene snp pair can act
- Each Tissue can be maximally loaded on 1 eigenconfig
- This means that it will be active only in tissues that are active in this eigenconfig

Thus each multivariate component will correspond to the importance of the direction of tissue 1 ...K

F_1