♠ Density and Parameters

- $\mathbf{x}$ is $p \times 1$ with mean vector $\mathbf{m}$ ($p \times 1$) and variance (covariance) matrix $\mathbf{V}$ ($p \times p$)
- Precision (or concentration) matrix $\mathbf{K} = \mathbf{V}^{-1}$ (assume non-singular)
- Density function

$$p(x) = c \exp(-(\mathbf{x}-\mathbf{m})'\mathbf{K}(\mathbf{x}-\mathbf{m})/2)$$

  with $c = |2\pi|^{p/2}|\mathbf{K}|^{1/2}$
- $\mathbf{x} \sim N(\mathbf{m}, \mathbf{V})$ or $\mathbf{x} \sim N(\mathbf{x}|\mathbf{m}, \mathbf{V})$

♠ Linear Transforms

- Any $k \times p$ matrix $\mathbf{G}$ and constant $k-$vector $\mathbf{a}$, $\mathbf{y} = \mathbf{a} + \mathbf{G}\mathbf{x}$ is normal $\mathbf{y} \sim N(\mathbf{a} + \mathbf{G}\mathbf{m}, \mathbf{G}\mathbf{V}\mathbf{G}')$
- $k < p$: Dimension reduction
- $k > p$: Rank deficient (singular) distribution

♠ Key Properties: Marginal & Conditional Distributions

Partition $\mathbf{x}$ as $\mathbf{x}_1$ and $\mathbf{x}_2$ and conformably partition $\mathbf{m}$ and $\mathbf{V}$ so that

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}, \qquad \mathbf{m} = \begin{pmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \end{pmatrix} \quad \& \quad \mathbf{V} = \begin{pmatrix} \mathbf{V}_1 & \mathbf{R} \\ \mathbf{R}' & \mathbf{V}_2 \end{pmatrix}$$

where $C(\mathbf{x}_1, \mathbf{x}_2) = \mathbf{R}$ (and of course $C(\mathbf{x}_2, \mathbf{x}_1) = \mathbf{R}'$.) Dimensions are conformable − any subsetting of $\mathbf{x}$ works

- $\mathbf{x}_1 \sim N(\mathbf{m}_1, \mathbf{V}_1)$ and $\mathbf{x}_2 \sim N(\mathbf{m}_2, \mathbf{V}_2)$
- Really critical to understanding regression are the conditional distributions: Here is $p(\mathbf{x}_1|\mathbf{x}_2)$ and the same general theory tells you what $p(\mathbf{x}_2|\mathbf{x}_1)$ is

$$(\mathbf{x}_1|\mathbf{x}_2) \sim N(\mathbf{a}_1 + \mathbf{B}_1\mathbf{x}_2, \mathbf{W}_1)$$

with

$$\mathbf{a}_1 = \mathbf{m}_1 - \mathbf{B}_1\mathbf{m}_2, \quad \mathbf{B}_1 = \mathbf{R}\mathbf{V}_2^{-1} \quad \& \quad \mathbf{W}_1 = \mathbf{V}_1 - \mathbf{B}_1\mathbf{R}'$$

♠ Precision Matrix and Dependencies

Take $\mathbf{x}_1 = x_1$, the first element of $\mathbf{x}$ so that $\mathbf{x}_2$ is all the rest. Another way of writing the conditional distribution above is in terms of the elements of the precision matrix $\mathbf{K}$ instead of $\mathbf{V}$ as follows (this is just based on standard linear algebra and representations of inverses of partitioned matrices).

- If $\mathbf{x}_1 = x_1$, then $\mathbf{B}_1$ is the $(p-1)$ row vector with j^{th} element

$$b_{1,j} = -K_{1,j}/K_{1,1}$$

  and $\mathbf{W}_1$ is the scalar variance $1/K_{1,1}$
- Shows the linear regression of x_1 (or any other x_i) on all other variables (genes)
- Note: Zeros in precision matrices corresponding to *conditional independencies*
- Underlies the major area of *Gaussian graphical models*

♠ Singular Normal

- $\mathbf{V}$ is singular; distribution is singular
- rank deficient: $rank(\mathbf{V}) = k < p$ – for some $k \times p$ matrix $\mathbf{G}$, $\mathbf{y} = \mathbf{G}\mathbf{x}$ has a non-singular distribution: variance matrix $\mathbf{G}\mathbf{V}\mathbf{G}'$ is non-singular.
- constrained linear combinations of $p - k$ elements of $\mathbf{x}$ – only k "free" dimensions
- density still has same form in terms of $\mathbf{K}$ where now $\mathbf{K} = \mathbf{V}^{-}$ is a generalised inverse of $\mathbf{V}$ (i.e., such that $\mathbf{K}\mathbf{V}\mathbf{K} = \mathbf{K}$ and $\mathbf{V}\mathbf{K}\mathbf{V} = \mathbf{V}$