Linear Algebra for Regression

- Cauchy Schwarz
- Projection
- Frisch-Waugh-Lovell Theorem } separate files.
- Variance Matrices – Precision Matrices
- Spectral Decomposition
- Matrix Factorization
- Geometry of the Bivariate Normal
- Schur Complement and Slicing Ellipsoids.
- Shadow / Slice = Marginal / Conditional
- Woodbury Formula : rank-k updates

Cauchy-Schwartz theorem

Let $\underset{\sim}{x}, \underset{\sim}{y} \in \mathbb{R}^n$

then $(\underset{\sim}{x}'\underset{\sim}{y})^2 \leq (\underset{\sim}{x}'\underset{\sim}{x})(\underset{\sim}{y}'\underset{\sim}{y})$

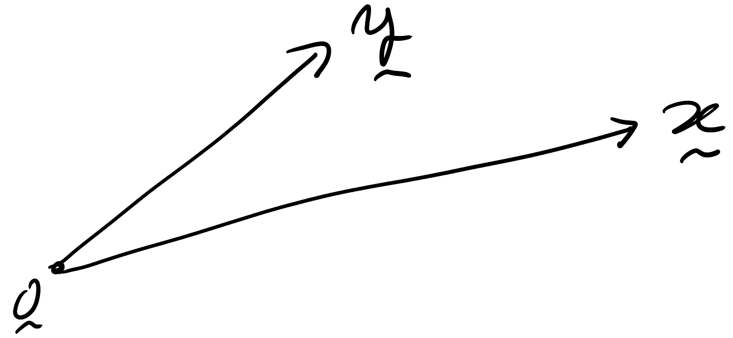
with equality iff $\exists\, a, b \in \mathbb{R}$, not both 0

$\ni a\underset{\sim}{x} + b\underset{\sim}{y} = \underset{\sim}{0}$

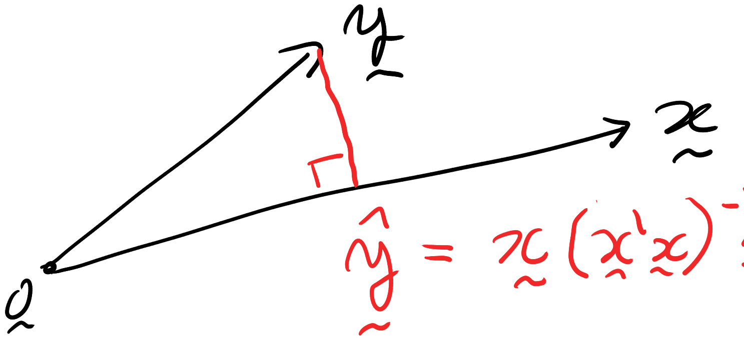
i.e. $\underset{\sim}{x}$ and $\underset{\sim}{y}$ are collinear with $\underset{\sim}{0}$

Note: $\underset{\sim}{x}$ and/or $\underset{\sim}{y}$ could be $= \underset{\sim}{0}$

Projections:



Projections:



$$\hat{\underset{\sim}{y}} = \underset{\sim}{x}(\underset{\sim}{x}'\underset{\sim}{x})^{-1}\underset{\sim}{x}'\underset{\sim}{y} = \underbrace{[\underset{\sim}{x}(\underset{\sim}{x}'\underset{\sim}{x})^{-1}\underset{\sim}{x}']}_{n \times n}\underbrace{\underset{\sim}{y}}_{n\times 1}$$

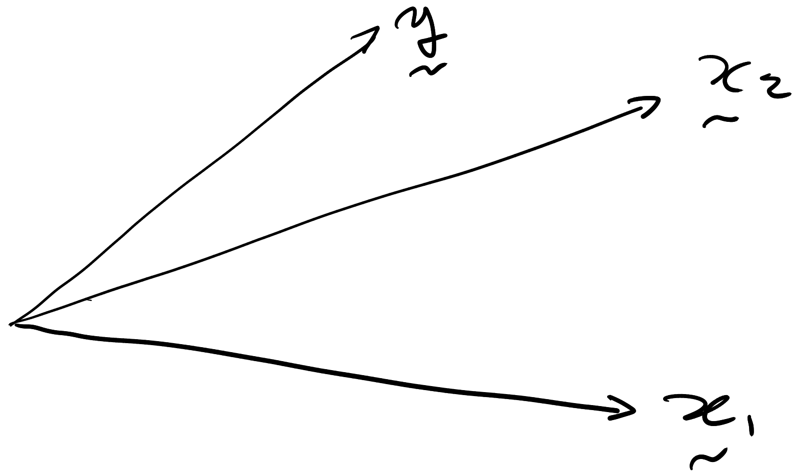
$$= H\underset{\sim}{y}$$

"hat matrix"

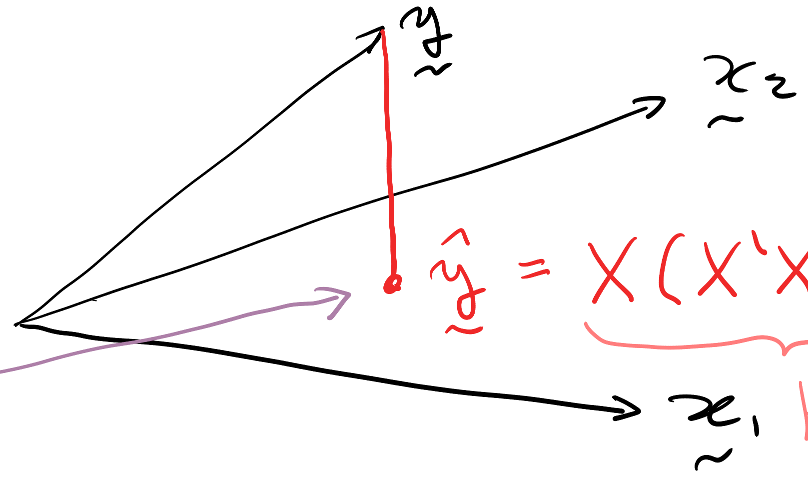
EX 1

Use the C-S inequality to show that $\hat{\underset{\sim}{y}}$ is the point in span$(\underset{\sim}{x})$ that is closest to $\underset{\sim}{y}$

With 2 $\underset{\sim}{x}$'s in $\mathbb{R}^n$



With 2 $\underset{\sim}{x}$'s in $\mathbb{R}^n$



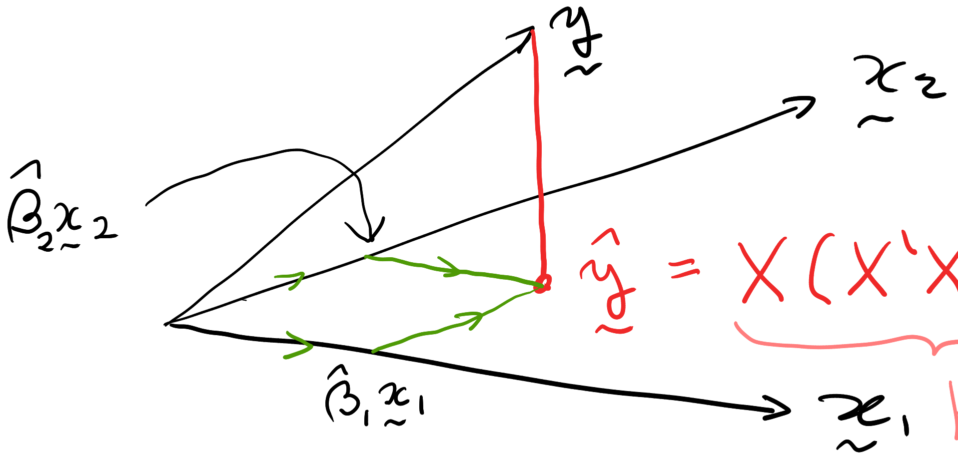
$$\underset{\sim}{\hat{y}} = \underbrace{X(X'X)^{-1}X'}_{H}\underset{\sim}{y} \quad \text{where } X = [\underset{\sim}{x}_1, \underset{\sim}{x}_2]$$

orthogonal projection of y onto span$(\underset{\sim}{x}_1, \underset{\sim}{x}_2)$

$$= \left\{ \underset{\sim}{x} : \underset{\sim}{x} = \beta_1 \underset{\sim}{x}_1 + \beta_2 \underset{\sim}{x}_2 \text{ for some } \begin{pmatrix}\beta_1 \\ \beta_2\end{pmatrix} \in \mathbb{R}^2 \right\}$$

$$= \left\{ X\underset{\sim}{\beta} : \underset{\sim}{\beta} \in \mathbb{R}^2 \right\}$$

With 2 $\underset{\sim}{x}$'s in $\mathbb{R}^n$ ($n=3$ here)

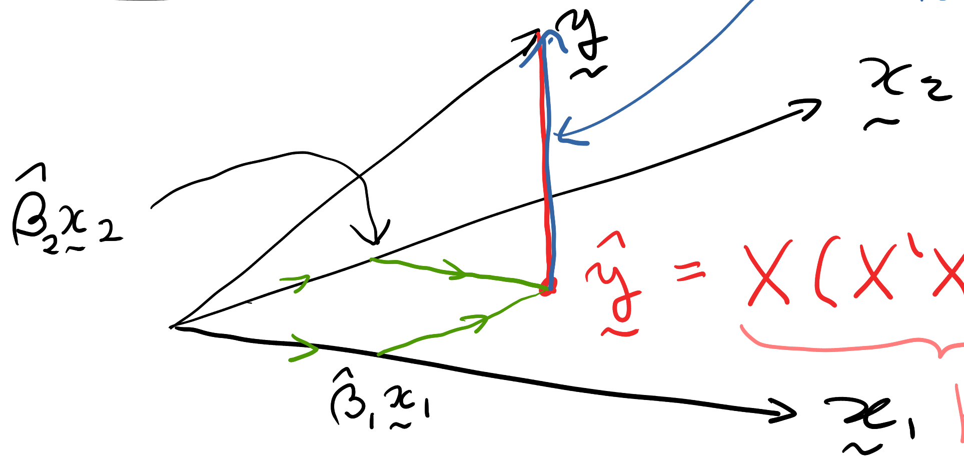


$$\hat{\underset{\sim}{y}} = \underbrace{X(X'X)^{-1}X'}_{H}\,\underset{\sim}{y} \quad \text{where } X = [\underset{\sim}{x}_1, \underset{\sim}{x}_2]$$

$$\hat{\underset{\sim}{y}} = \underset{\sim}{x}_1 \hat{\beta}_1 + \underset{\sim}{x}_2 \hat{\beta}_2 = \begin{bmatrix} \underset{\sim}{x}_1 & \underset{\sim}{x}_2 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = X \hat{\underset{\sim}{\beta}} \quad \text{where } \hat{\underset{\sim}{\beta}} = (X'X)^{-1}X'\underset{\sim}{y}$$

With 2 $\underset{\sim}{x}$'s in $\mathbb{R}^n$

$$\underset{\sim}{e} = \underset{\sim}{y} - \hat{\underset{\sim}{y}} = (I - H)\underset{\sim}{y}$$



$$\hat{\underset{\sim}{y}} = \underbrace{X(X'X)^{-1}X'}_{H}\underset{\sim}{y} \quad \text{where } X = [\underset{\sim}{x}_1, \underset{\sim}{x}_2]$$

$$\hat{\underset{\sim}{y}} = \underset{\sim}{x}_1 \hat{\beta}_1 + x_2 \hat{\beta}_2 = \begin{bmatrix} \underset{\sim}{x}_1 & \underset{\sim}{x}_2 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = X \hat{\underset{\sim}{\beta}} \quad \text{where } \hat{\underset{\sim}{\beta}} = (X'X)^{-1}X'\underset{\sim}{y}$$

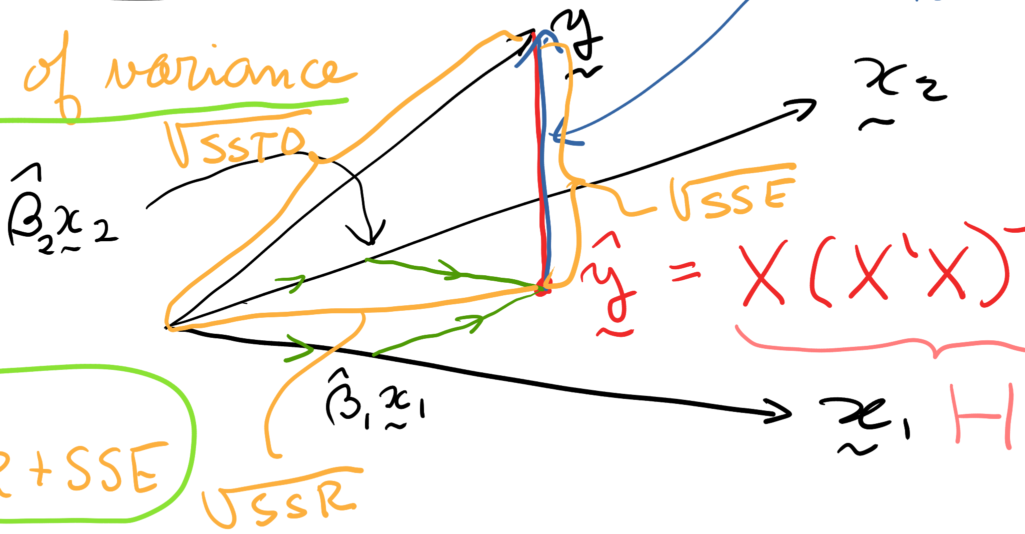
Note: $H = H^2$ ($\therefore$ projection matrix) and $H = H'$ ($\therefore$ orthogonal projection)

$$(I - H) = (I - H)^2 \qquad\qquad (I - H) = (I - H)'$$

With 2 $\underset{\sim}{x}$'s in $\mathbb{R}^n$

Analysis of variance

$$\underset{\sim}{e} = \underset{\sim}{y} - \hat{\underset{\sim}{y}} = (I - H)\underset{\sim}{y}$$



$$\hat{\underset{\sim}{y}} = \underbrace{X(X'X)^{-1}X'}_{H}\underset{\sim}{y} \quad \text{where } X = [\underset{\sim}{x}_1, \underset{\sim}{x}_2]$$

$$SSTO = SSR + SSE$$

$$\hat{\underset{\sim}{y}} = \underset{\sim}{x}_1 \hat{\beta}_1 + \underset{\sim}{x}_2 \hat{\beta}_2 = \begin{bmatrix} \underset{\sim}{x}_1 & \underset{\sim}{x}_2 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = X \hat{\underset{\sim}{\beta}} \quad \text{where } \hat{\underset{\sim}{\beta}} = (X'X)^{-1}X'\underset{\sim}{y}$$

Note: $H = H^2$ ($\therefore$ projection matrix) and $H = H'$ ($\therefore$ orthogonal projection)

$(I-H) = (I-H)^2$ ″ $\qquad (I-H) = (I-H)'$ ″

Statistical Model:

Def $\boxed{\underset{n\times 1}{\underset{\sim}{Y}} = \underset{n\times p}{X}\,\underset{p\times 1}{\underset{\sim}{\beta}} + \underset{n\times 1}{\underset{\sim}{\varepsilon}}}$

$\underset{\sim}{Y}$, X observed, $\underset{\sim}{\beta}$ unknown, $\underset{\sim}{\varepsilon}$ unobserved,

σ^2 known or unknown

Normal regression model

Def $\boxed{\underset{\sim}{\varepsilon} \sim N(\underset{\sim}{0}, \sigma^2 I)}$ $\quad \hat{\underset{\sim}{\beta}} = (X'X)^{-1}X'\underset{\sim}{Y}$ is UMVUE of $\underset{\sim}{\beta}$

Unif. Min. Var. Unb. Estimator

"BUE" Best Unbiased Estimator among

"General" model

Def. $\boxed{\underset{\sim}{\varepsilon} \sim ?(\underset{\sim}{0}, \sigma^2 I)}$ $\quad \hat{\underset{\sim}{\beta}} = (X'X)^{-1}X'\underset{\sim}{Y}$ is BLUE of $\underset{\sim}{\beta}$

any mean variance

Best Linear Unbiased Estimator among all

Gauss-Markov Theorem

$$\mathcal{E}_r = \{ \underset{\sim}{\beta} : (\underset{\sim}{\beta} - \underset{\sim}{\beta_0})' \underbrace{(\sigma^2 (X'X)^{-1})^{-1}}_{\left(\frac{1}{\sigma^2} X'X\right)} (\underset{\sim}{\beta} - \underset{\sim}{\beta_0}) = r^2 \}$$

If σ^2 unknown can use

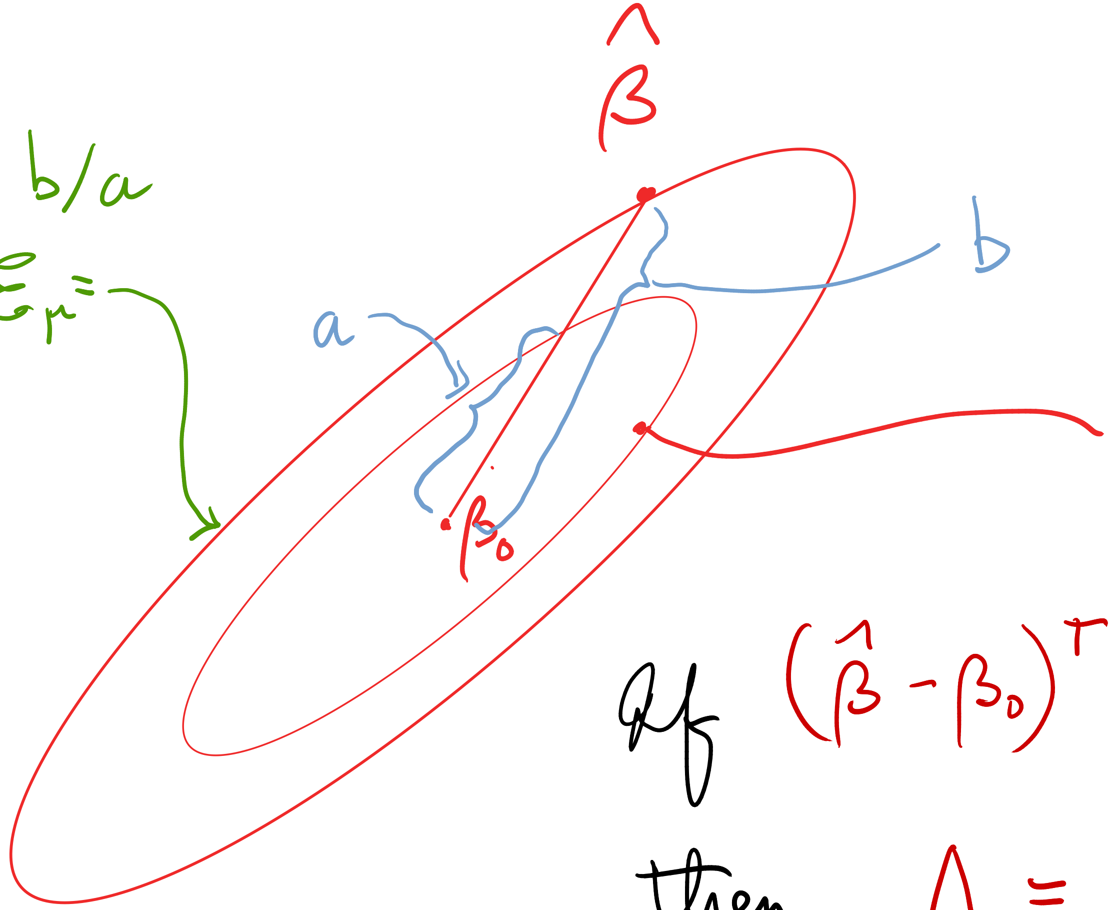
MLE: $\hat{\sigma}^2 = \hat{\sigma}_\varepsilon^2 = \dfrac{SSE}{n}$ $\qquad E(\hat{\sigma}^2) = \left(\dfrac{n-p}{n}\right)\sigma^2$

Unbiased Estimate: $S^2 = S_\varepsilon^2 = \dfrac{SSE}{n-p}$ $\qquad E(S^2) = \sigma^2$

If $n >> p$ $\qquad E(\hat{\sigma}^2) \approx E(S^2)$

Let $r = b/a$

then $\mathcal{E}_r =$



$\{\underset{\sim}{\beta} : \underset{\sim}{\beta}' V^{-1} \underset{\sim}{\beta} = 1\}$

= "unit" concentration ellipse

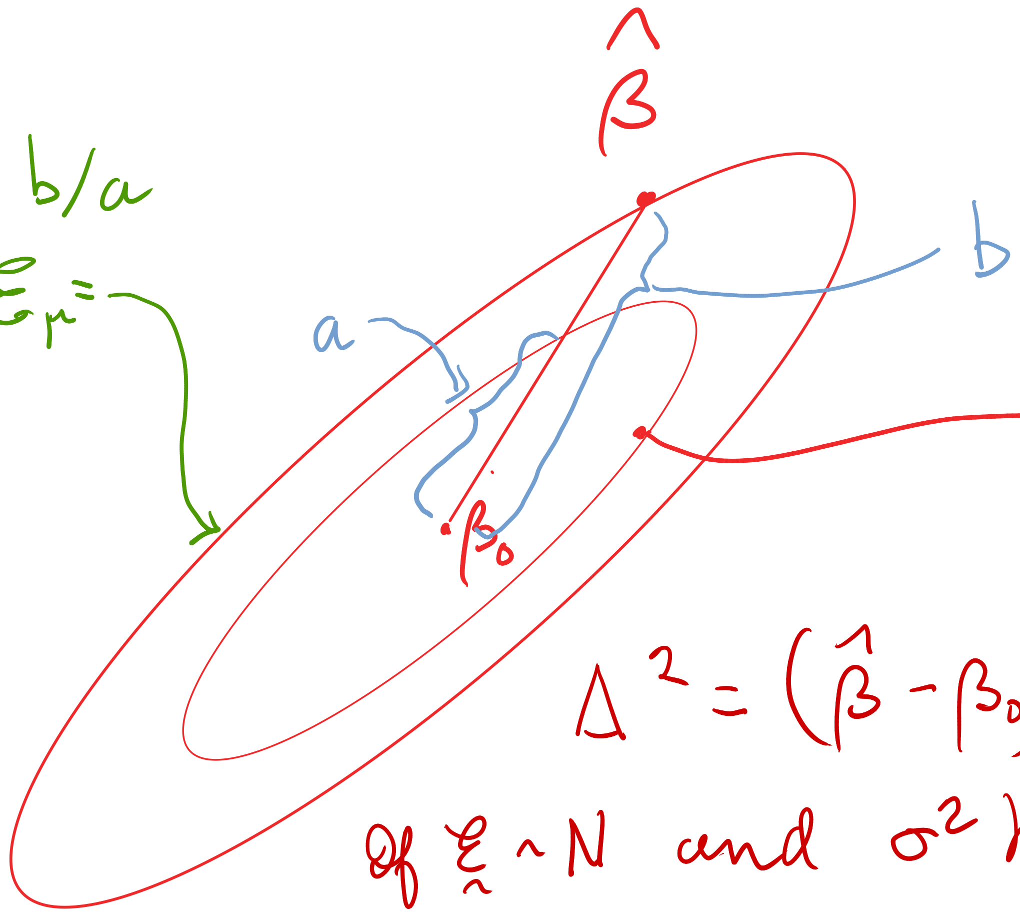
If $(\hat{\beta} - \beta_0)^T V^{-1} (\hat{\beta} - \beta_0) = \Delta^2$

then $\Delta = b/a$

Δ is "statistical distance of $\hat{\beta}$ from $E(\hat{\underset{\sim}{\beta}}_\gamma) = \underset{\sim}{\beta}_2$

Let $r = b/a$

then $\mathcal{E}_\mu =$



$\{\underset{\sim}{\beta} : \hat{\underset{\sim}{\beta}}' X'X \underset{\sim}{\beta} = 1\}$

= "unit" concentration ellipse

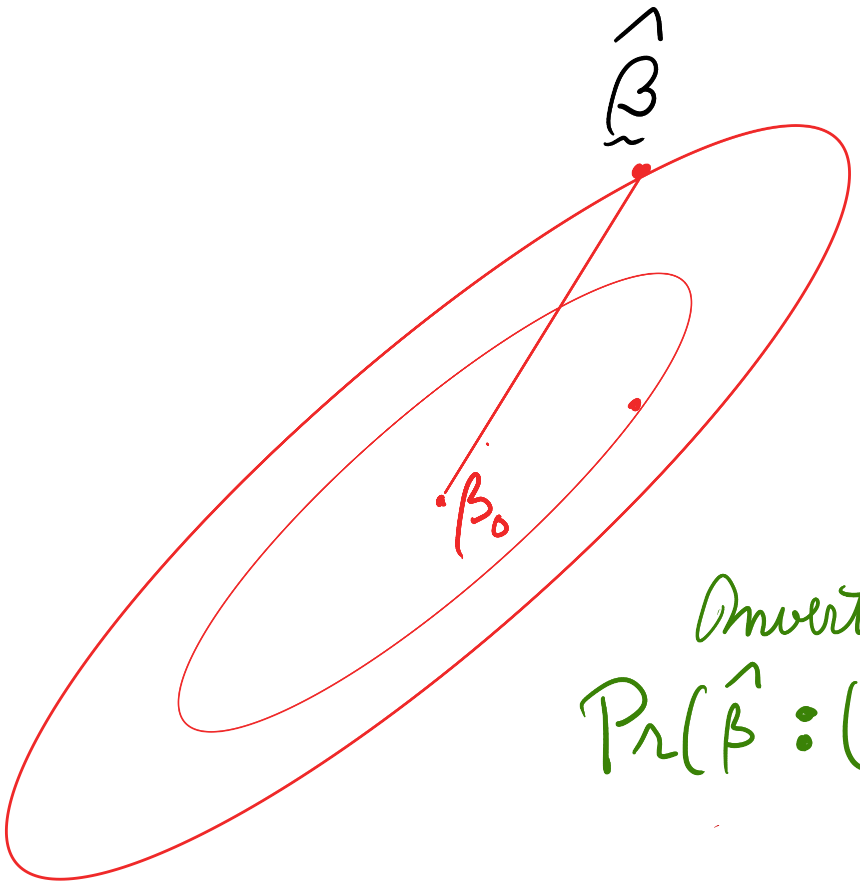
$$\Delta^2 = (\hat\beta - \beta_0)^T X'X (\hat\beta - \beta_0)$$

If $\underset{\sim}{\epsilon} \sim N$ and σ^2 known then $\frac{\Delta}{\sigma} \sim \sqrt{\chi^2_d}$

" " $s^2 = \frac{SSE}{n-p}$ then $\frac{\Delta}{s} \sim \sqrt{d F_{d,\nu}}$

$\nu = n - p$

Confidence Regions & Tests

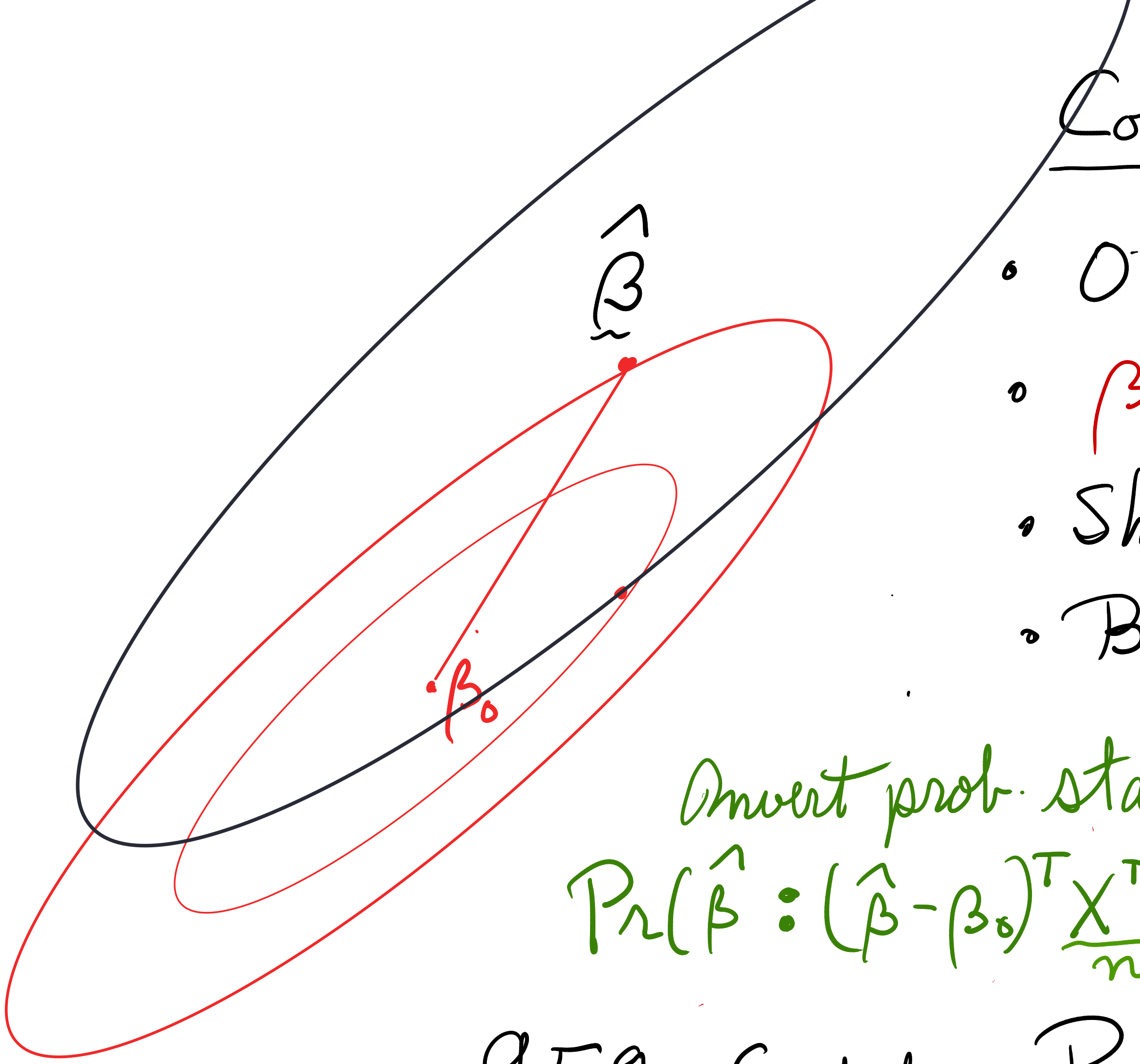


- Observe $\hat{\beta}$
- β_0 unknown
- Shape of ellipse known
- But not radius

Invert prob. statement:

$$\Pr\left(\hat{\beta} : (\hat{\beta}-\beta_0)^T X^T X (\hat{\beta}-\beta_0) \le s^2 \times d F_{d,\nu}^{.95}\right) = 0.95$$

Confidence Regions & Tests



- Observe $\hat\beta$
- β_0 unknown
- Shape of ellipse known
- But not radius

Invert prob. statement:

$$\Pr\left(\hat\beta : (\hat\beta-\beta_0)^T \frac{X^TX}{n}(\hat\beta-\beta_0) \le \frac{s^2}{n} \times d F^{.95}_{d,\nu}\right) = 0.95$$

95% Confidence Region

$$= \left\{ \underset{\sim}{\beta} : (\beta-\hat\beta)^T \frac{X^TX}{n} (\beta-\hat\beta) \le \frac{s^2}{n} d F^{.95}_{d,\nu} \right\}$$

Confidence Regions & Tests

95% CE $= \mathcal{E}^{.95}$

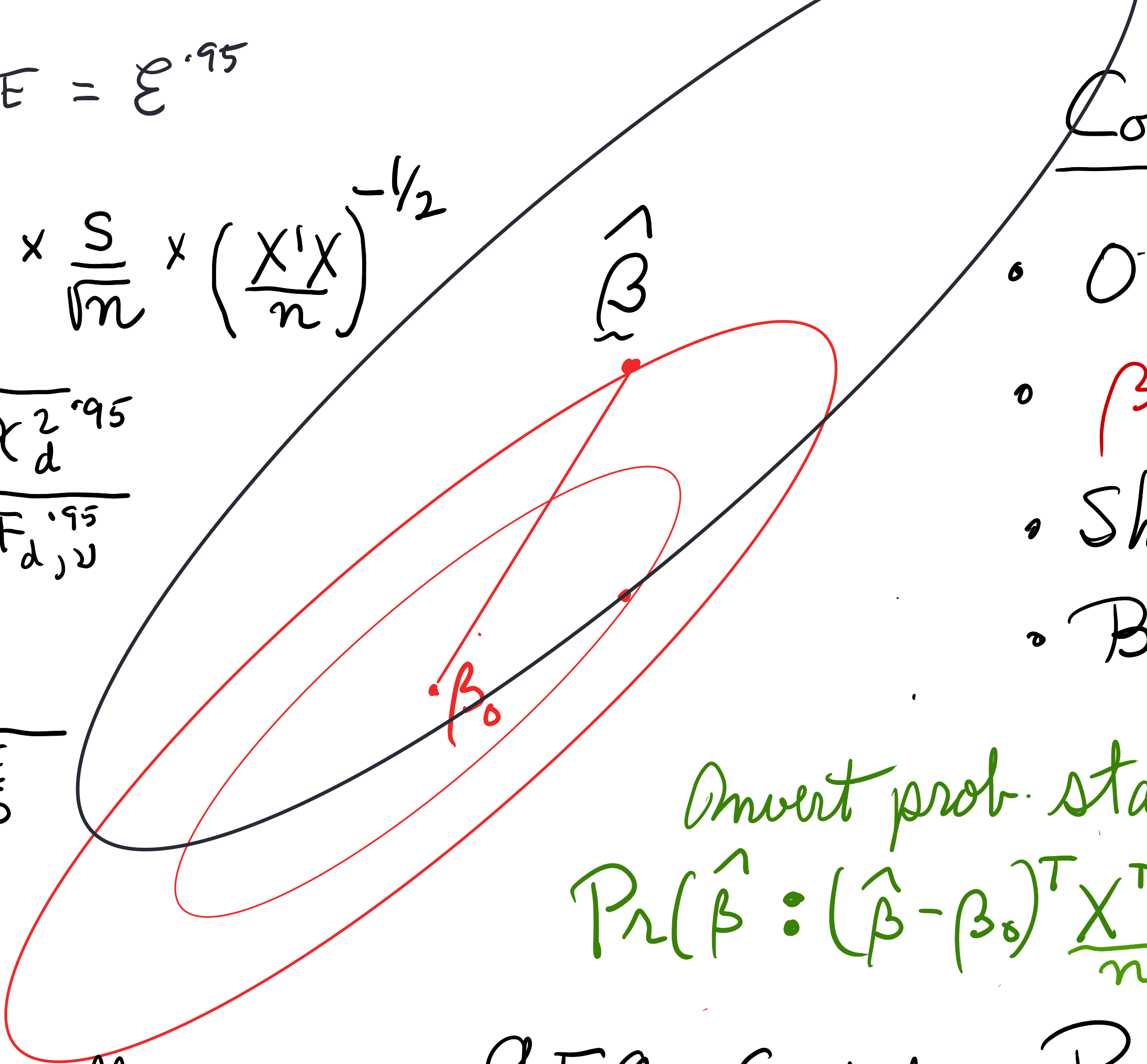
$$= \hat{\beta} \oplus D \times \frac{S}{\sqrt{n}} \times \left(\frac{X'X}{n}\right)^{-1/2}$$

$$D = \begin{cases} \sqrt{\chi^{2\,.95}_{d}} \\ \sqrt{d\,F^{.95}_{d,\nu}} \end{cases}$$

$$S = \begin{cases} \sigma \\ \sqrt{\frac{SSE}{n-p}} \end{cases}$$

$\sqrt{\frac{X'X}{n}}^{-1}$ denotes the ellipse

$$\mathcal{E} = \left\{ \underset{\sim}{u} : \underset{\sim}{u}'\left(\frac{X'X}{n}\right)u = 1 \right\}$$



- Observe $\hat{\beta}$
- β_0 unknown
- Shape of ellipse known
- But not radius

Invert prob. statement:

$$\Pr\left(\hat{\beta} : (\hat{\beta}-\beta_0)^T \frac{X^T X}{n} (\hat{\beta}-\beta_0) \le \frac{S^2}{n} \times d F^{.95}_{d,\nu}\right) = 0.95$$

95% Confidence Region

$$= \left\{ \underset{\sim}{\beta} : (\beta-\hat{\beta})^T \frac{X^T X}{n} (\beta-\hat{\beta}) \le \frac{S^2}{n} d F^{.95}_{d,\nu} \right\}$$

Confidence Regions & Tests

95% CE $= \mathcal{E}^{.95}$

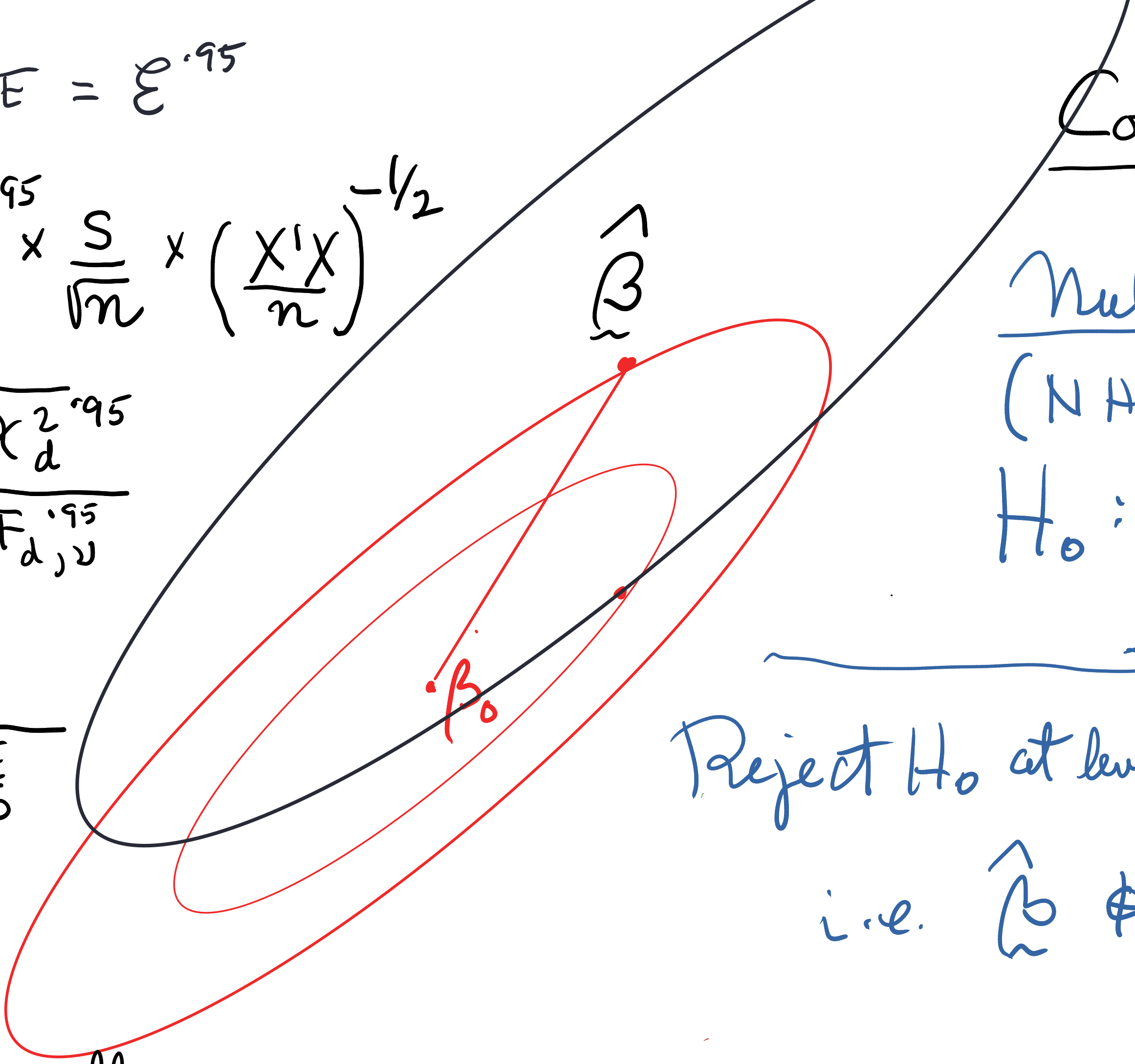
$$= \hat{\beta} \oplus D^{.95} \times \frac{S}{\sqrt{n}} \times \left(\frac{X'X}{n}\right)^{-1/2}$$

$$D^{.95} = \begin{cases} \sqrt{\chi^{2\;.95}_{d}} \\ \sqrt{d\,F^{.95}_{d,\nu}} \end{cases}$$

$$S = \begin{cases} \sigma \\ \sqrt{\frac{SSE}{n-p}} \end{cases}$$

$\sqrt{\frac{X'X}{n}}^{\,-1}$ denotes the ellipse

$$\mathcal{E} = \left\{ \underset{\sim}{u} : \underset{\sim}{u}'\left(\frac{X'X}{n}\right)\underset{\sim}{u} = 1 \right\}$$



Null Hypothesis Significance Test (NHST)

$H_0 : \underset{\sim}{\beta} = \underset{\sim}{\beta}_0$ vs. $H_A : \underset{\sim}{\beta} \neq \underset{\sim}{\beta}_0$

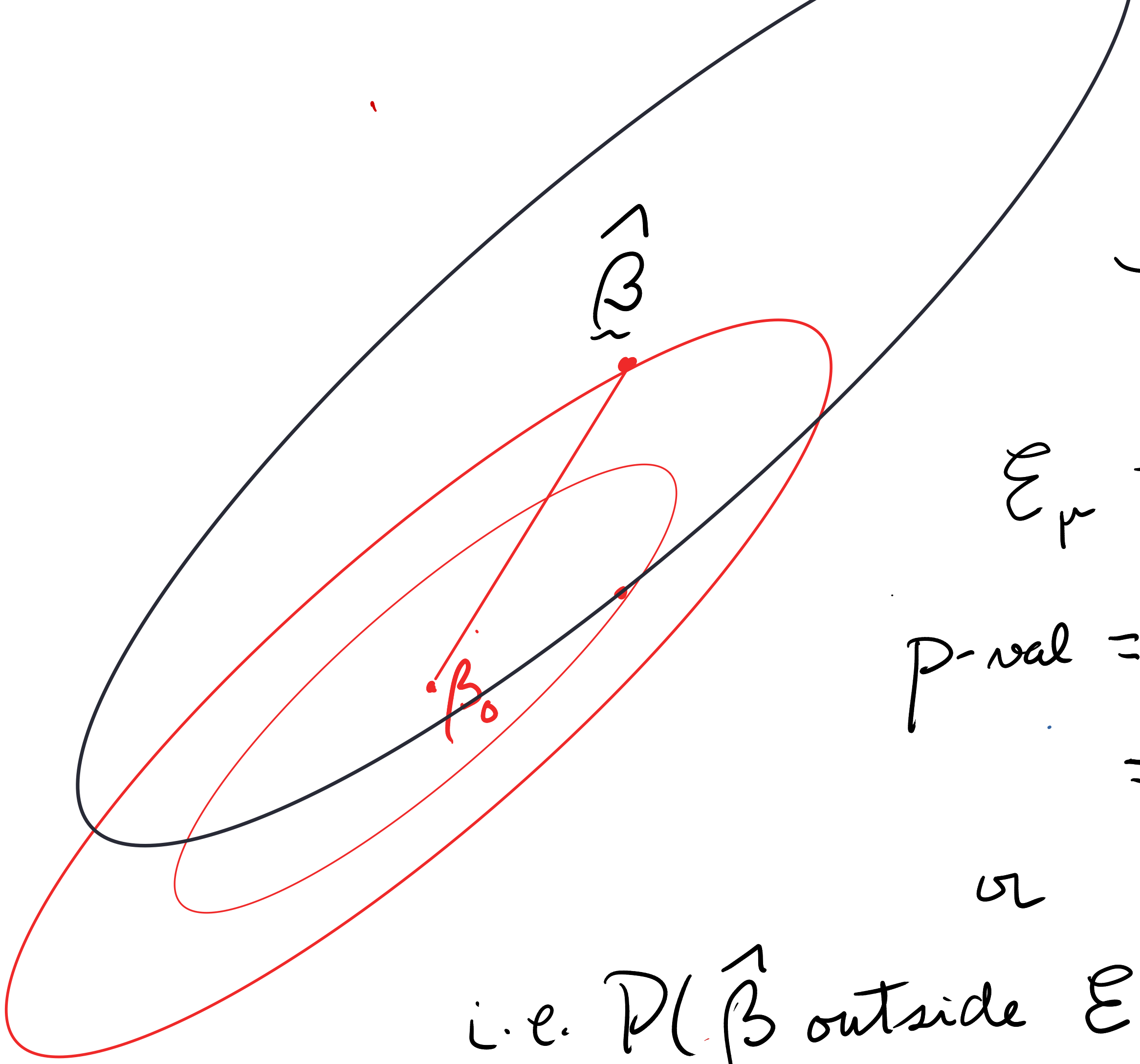
true

Reject H_0 at level α if $\hat{\beta} \notin$ Acceptance region

i.e. $\hat{\underset{\sim}{\beta}} \notin \underset{\sim}{\beta}_0 \oplus D^{.95} \frac{S}{\sqrt{n}} \left(\frac{X'X}{n}\right)^{-1/2}$

equivalently: $\underset{\sim}{\beta}_0 \notin$ 95% Confidence ellipse

$$\hat{\underset{\sim}{\beta}} \oplus D^{.95} \frac{S}{\sqrt{n}} \left(\frac{X'X}{n}\right)^{-1/2}$$



p-values

Let $\hat{\underset{\sim}{\beta}} \in \mathcal{E}_\mu$

$$\mathcal{E}_\mu = \underset{\sim}{\beta}_0 + \mu \frac{s}{\sqrt{n}} \left(\frac{X'X}{n}\right)^{-1/2}$$

$$\text{p-val} = \Pr(\Delta > \mu \mid H_0)$$

$$= \Pr\left(\sqrt{d F_{d,\nu}} > \mu\right)$$

or $$= \Pr\left(\sqrt{\chi^2_d} > \mu\right)$$

i.e. $P(\hat{\beta}$ outside $\mathcal{E}_\mu \mid H_0)$

Variance Matrices & Factorizations

Following are equivalent

① Σ is a variance matrix i.e. $\exists$ RV $\underset{\sim}{X} \ni \Sigma = \mathrm{Var}(\underset{\sim}{X})$

② Σ is non-negative definite i.e $\forall \underset{\sim}{x},\ \underset{\sim}{x}'\Sigma\underset{\sim}{x} \geq 0$

③ $\exists$ orthog. Γ, diag Λ with $\Lambda \geq 0$ Spectral decomposition

$$\ni \Sigma = \Gamma \Lambda \Gamma'$$

④ $\exists A \ni \Gamma = AA^T$

We can require A to be, e.g. +many others

- Lower triangular — Left Cholesky
- Upper triangular — Right Cholesky
- non-negative definite — "square root"
- orthogonal columns — principal components

$\boxed{\text{EX 2 : Prove}}$

non-singular

Variance Matrices & Factorizations

Following are equivalent

① Σ is a non-singular variance matrix i.e. $\exists$ RV $\underset{\sim}{X} \ni \Sigma = Var(\underset{\sim}{X})$

② Σ is ~~non-negative~~ positive definite i.e $\forall\ \underset{\sim}{x} \ni \underset{\sim}{x} \neq 0$, $\underset{\sim}{x}'\Sigma\underset{\sim}{x} \cancel{\geq} > 0$

③ $\exists$ orthog. Γ, diag Λ with $\Lambda \cancel{\geq} 0$ $\lambda_i > 0\ \forall i$ Spectral decomposition

$\ni \Sigma = \Gamma \Lambda \Gamma'$

④ $\exists\ A \ni \Gamma = AA^T$

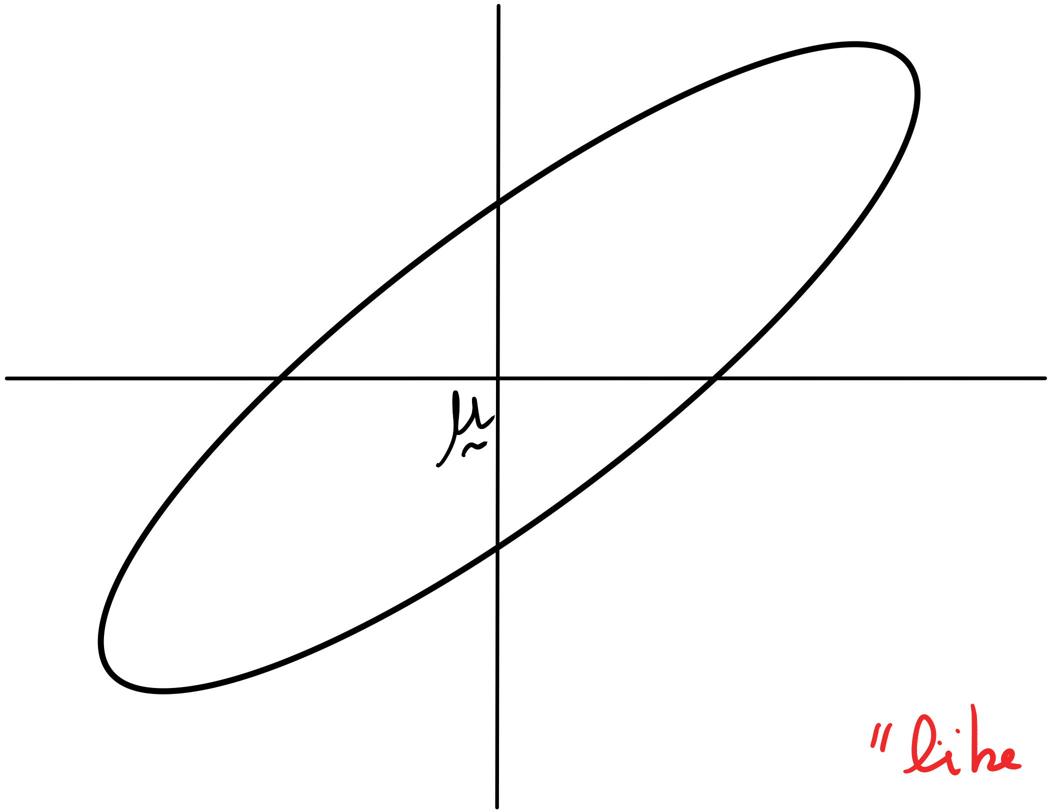
A non-singular

We can require A to be, e.g. + many others

- Lower triangular — Left Cholesky
- Upper triangular — Right Cholesky
- non-negative definite — "square root"
- orthogonal columns — principal components

$\boxed{\text{EX 3 Prove}}$

The Shape of Ellipses



Let $\underset{\sim}{X} \sim N(\underset{\sim}{\mu}, \Sigma)$

"Standard concentration ellipse"

$$f(\underset{\sim}{x}; \underset{\sim}{\mu}, \Sigma) = \frac{1}{(2\pi)^{p/2}} \frac{1}{|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(\underset{\sim}{x}-\underset{\sim}{\mu})' \Sigma^{1/2} (\underset{\sim}{x}-\underset{\sim}{\mu})\right\}$$

$$\mathcal{E} = \{\underset{\sim}{x} : \underset{\sim}{x}' \Sigma^{-1} \underset{\sim}{x} = 1\}$$

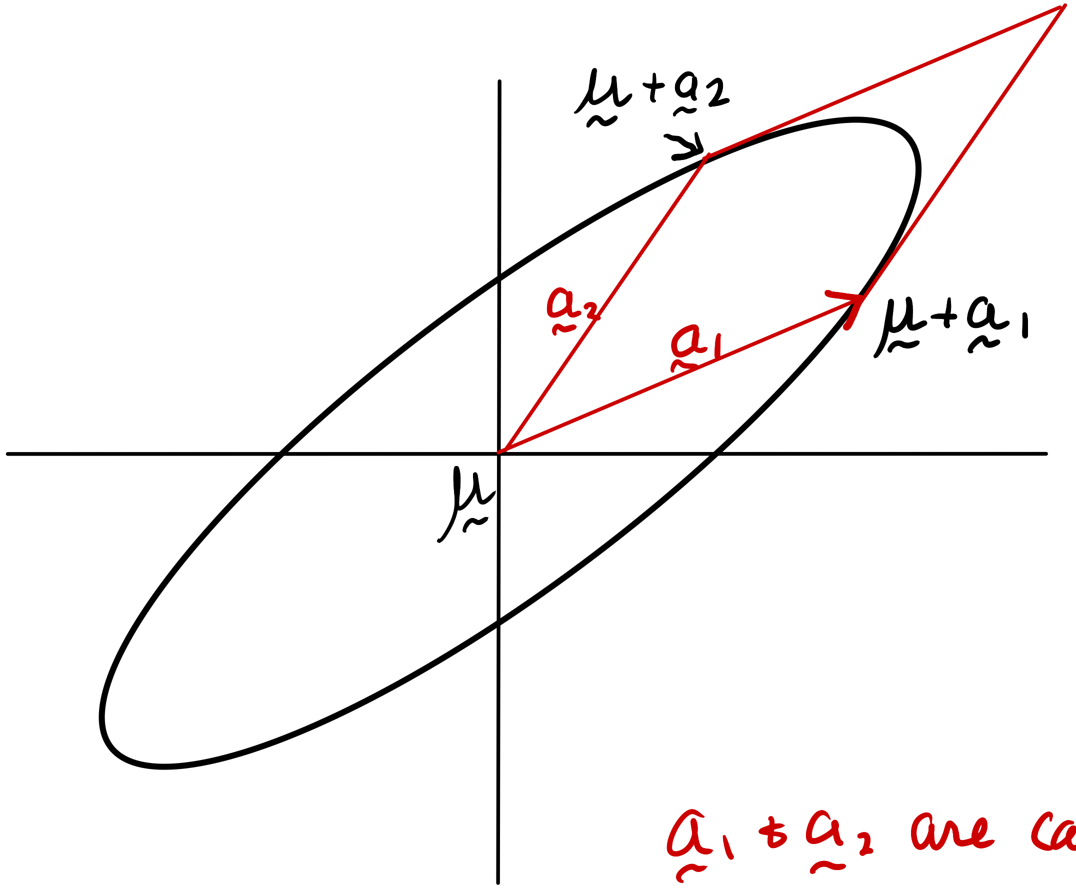
"standard concentration ellipse"

$$\mathcal{E} = \underset{\sim}{\mu} \oplus \sqrt{\Sigma}$$

"like a multivariate SD interval"

$\mu \pm \sigma$

The Shape of Ellipses



Let $\underset{2\times 2}{A}$ be any factor of $\underset{2\times 2}{\Sigma}$

i.e. $\Sigma = A\,A'$

Let $A = [\underset{\sim}{a}_1 \; \underset{\sim}{a}_2]$

Then

$$\underset{\sim}{\mu} + \underset{\sim}{a}_i \in \mathcal{E}$$

and the tangent to $\mathcal{E}$

at $\underset{\sim}{\mu}+\underset{\sim}{a}_i$ is $\parallel$ $\underset{\sim}{a}_{3-i}$

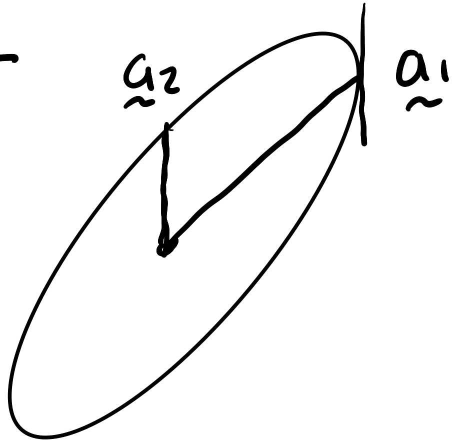
$\underset{\sim}{a}_1$ & $\underset{\sim}{a}_2$ are called conjugate radii of $\mathcal{E}$

Axes through $\underset{\sim}{a}_1, a_2$ are conjugate axes of $\mathcal{E}$

Factorizations

Left Choleski

$$L = \begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{a}_1 & \underset{\sim}{a}_2 \end{bmatrix}$$

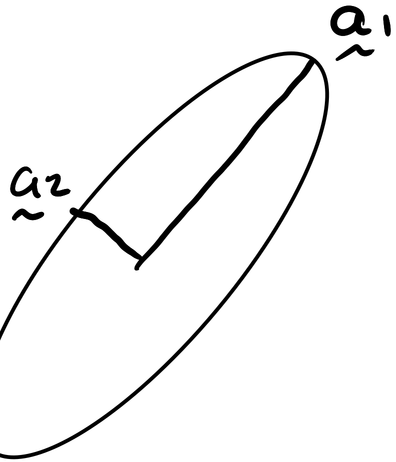


Principal Components

$$\Sigma = \Gamma \Lambda \Gamma'$$

$$= \underbrace{\Gamma \Lambda^{1/2}}_{A} \Lambda^{1/2} \Gamma'$$

$$= A \quad A'$$



etc.

EX3: What does a Right Cholesky factorization look like

Spectral Decomposition Theorem (SDT)

Let $\underset{n \times n}{A}$ be symmetric. Then $\exists$ orthogonal $\underset{n \times n}{\Gamma}$ and diagonal $\underset{n \times n}{\Lambda} = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$

$$\ni \quad A = \Gamma \Lambda \Gamma'$$

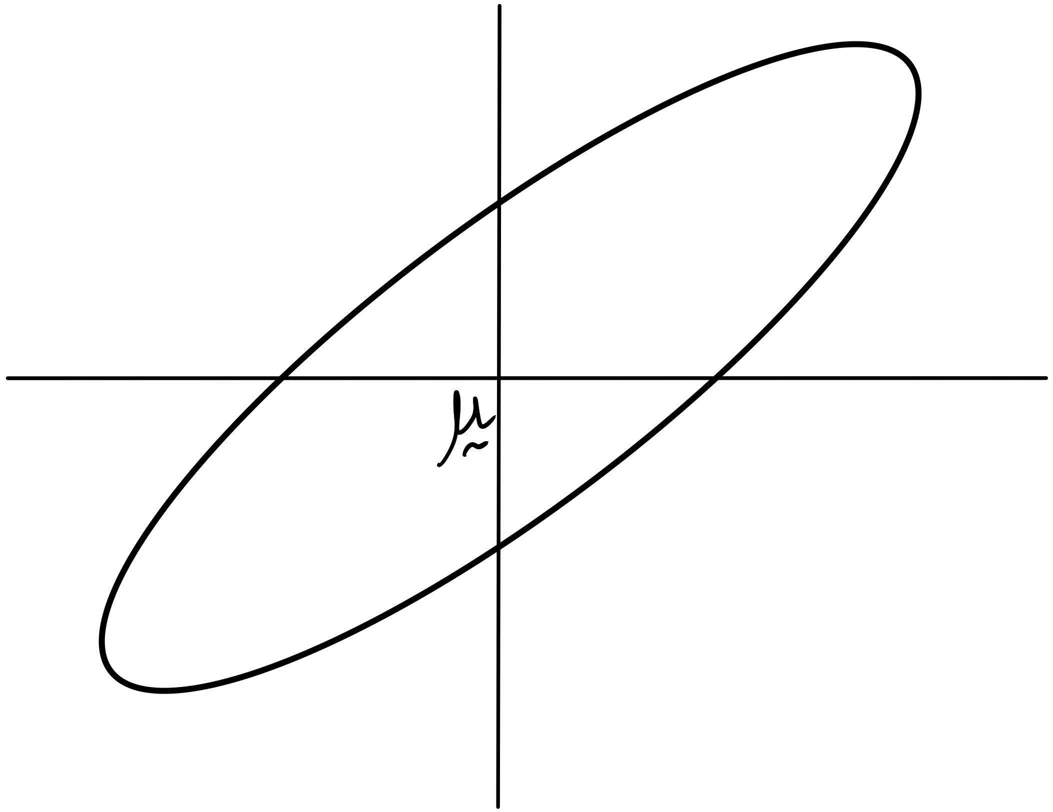
Note: Letting $\Gamma = [\underset{\sim}{\gamma}_1 \cdots \underset{\sim}{\gamma}_n]$

$\underset{\sim}{\gamma}_i$ is an eigenvector belonging to eigenvalue λ_i

i.e. $\quad A \underset{\sim}{\gamma}_i = \lambda_i \underset{\sim}{\gamma}_i$

$\boxed{\text{EX 4: Find 2 references to proofs}}$

The Shape of Ellipses



Let $\underset{\sim}{X} \sim N(\underset{\sim}{\mu}, \Sigma)$

"Standard concentration ellipse"

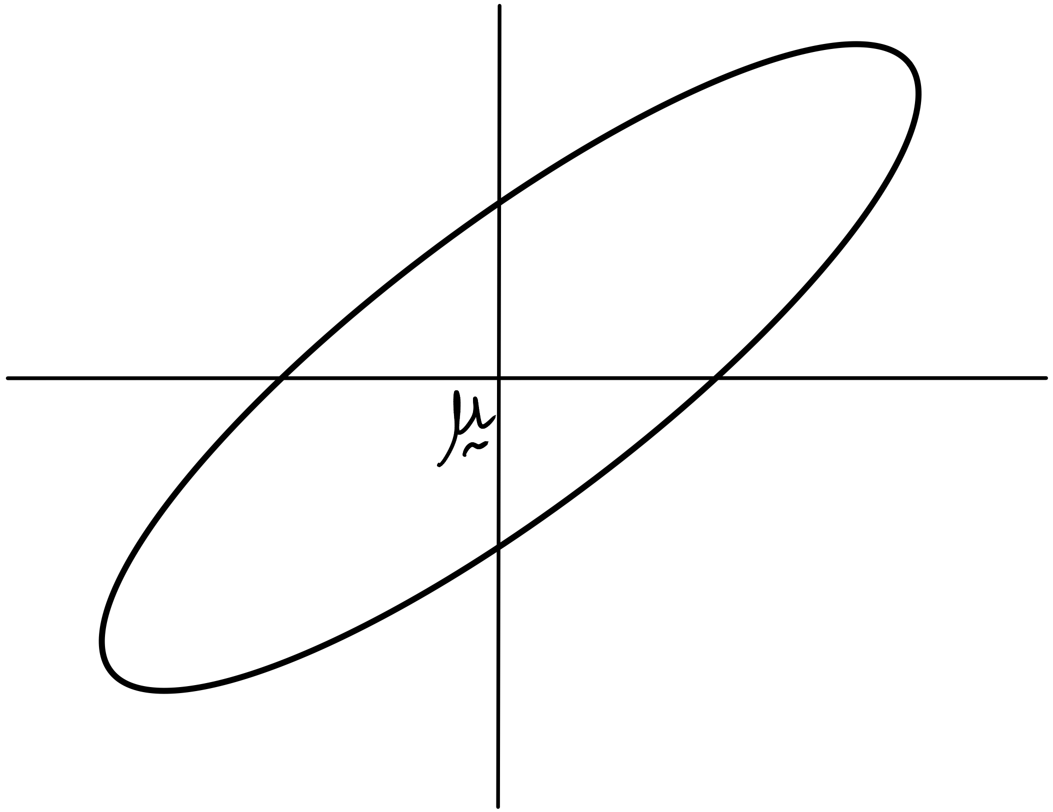
$$f(\underset{\sim}{x}; \underset{\sim}{\mu}, \Sigma) = \frac{1}{(2\pi)^{p/2}} \frac{1}{|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(\underset{\sim}{x}-\underset{\sim}{\mu})' \Sigma^{-1/2} (\underset{\sim}{x}-\underset{\sim}{\mu})\right\}$$

$$\mathcal{E} = \{\underset{\sim}{x} : \underset{\sim}{x}' \Sigma^{-1} \underset{\sim}{x} = 1\}$$

"standard concentration ellipse"

$$\mathcal{E} = \underset{\sim}{\mu} \oplus \sqrt{\Sigma}$$

The Shape of Ellipses



Whether N or ?

If $\underset{\sim}{Y} = T(\underset{\sim}{X})$

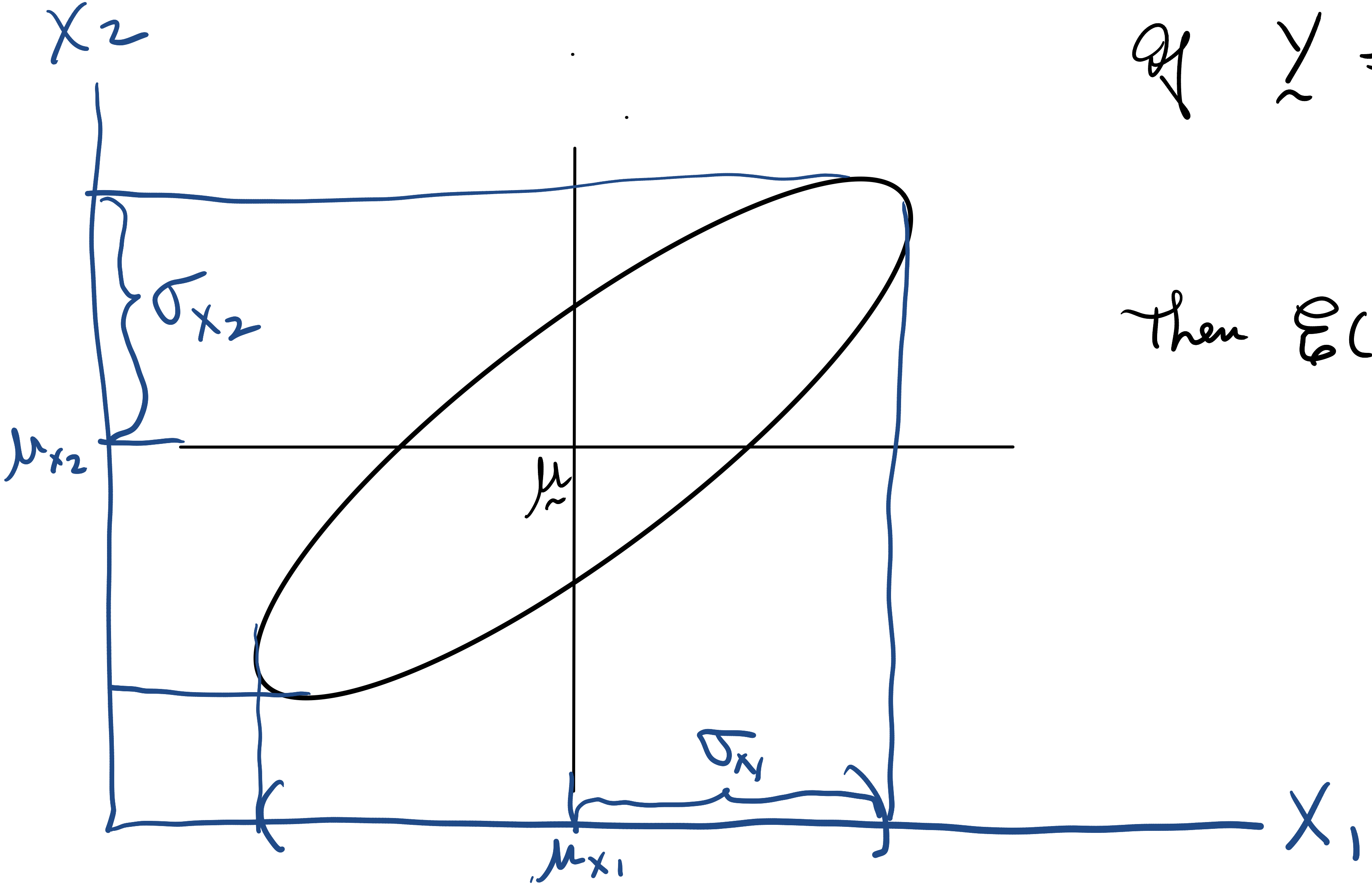
affine transformation

i.e. $\underset{\sim}{Y} = \underset{\sim}{a} + B\underset{\sim}{X}$

Then $\mathcal{E}(\underset{\sim}{Y}) = T\,\mathcal{E}(\underset{\sim}{X})$

$\boxed{\text{EX 5: Prove}}$

The Shape of Ellipses



Whether N or ?

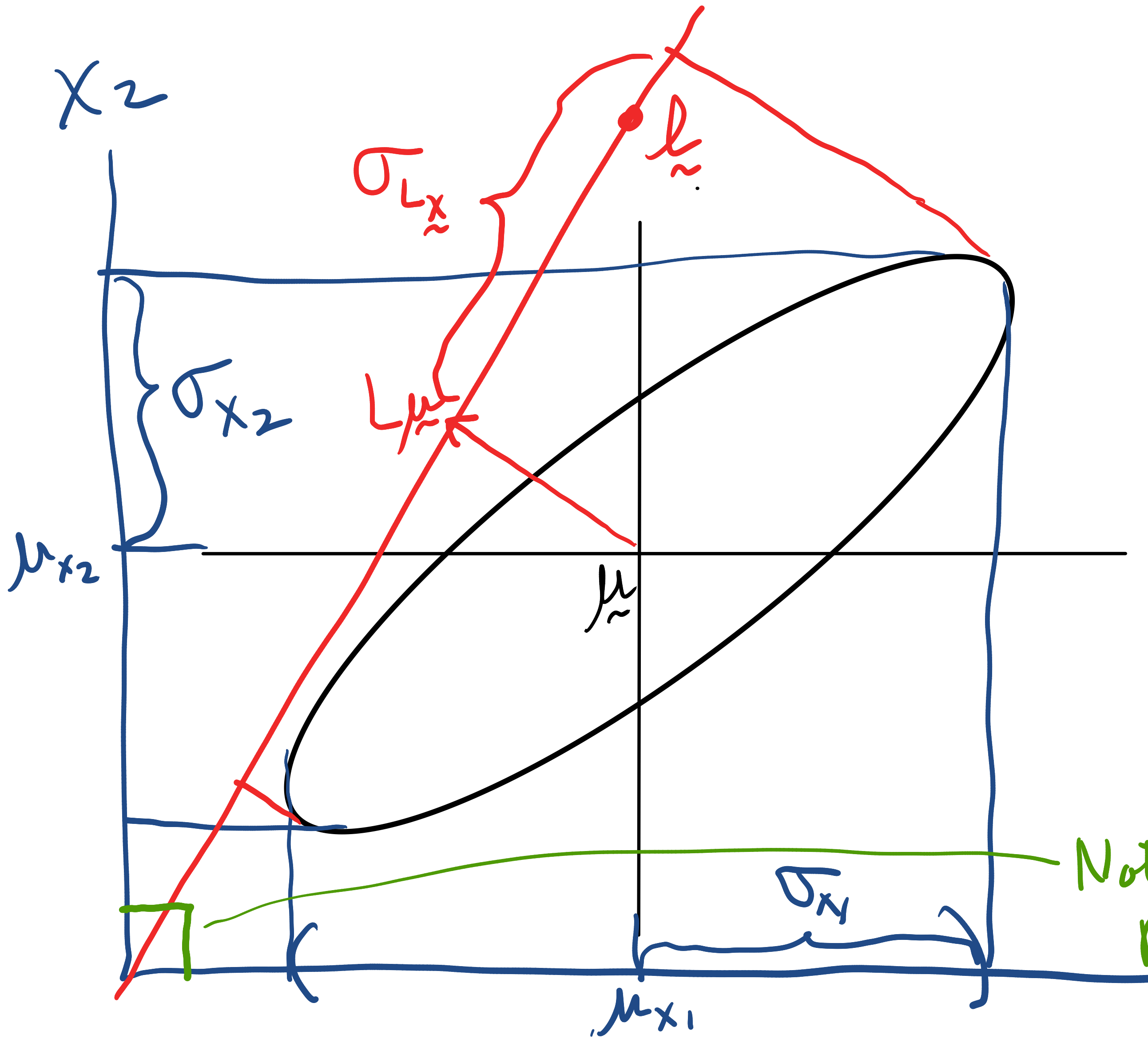
If $\underset{\sim}{Y} = T(\underset{\sim}{X})$

$\uparrow$ affine transformation

i.e. $\underset{\sim}{Y} = \underset{\sim}{a} + B\underset{\sim}{X}$

Then $\mathcal{E}(\underset{\sim}{Y}) = T\,\mathcal{E}(\underset{\sim}{X})$

The Shape of Ellipses



Whether N or ?

If $\underset{\sim}{Y} = T(\underset{\sim}{X})$

affine transformation

i.e. $\underset{\sim}{Y} = \underset{\sim}{a} + B\underset{\sim}{X}$

Then $\mathcal{E}(\underset{\sim}{Y}) = T\,\mathcal{E}(\underset{\sim}{X})$

Projection onto span($\underset{\sim}{\ell}$)

$L = \underset{\sim}{\ell}(\underset{\sim}{\ell}'\underset{\sim}{\ell})^{-1}\underset{\sim}{\ell}'$

same size

$Y = L\underset{\sim}{X}$

Note: Orthogonal projection looks perpendicular only if graph is "Euclidean" - units of X_1, X_2 are

E & Var under lin. transf.

$$\underset{\sim}{Y} = \underset{\sim}{a} + B\underset{\sim}{X}$$

$$E(\underset{\sim}{Y}) = \underset{\sim}{a} + B\,E(\underset{\sim}{X})$$

$$\mathrm{Var}(\underset{\sim}{Y}) = B\,\mathrm{Var}(\underset{\sim}{X})\,B'$$

$$\boxed{\text{EX 6 Prove}}$$

Mult. normal : $\begin{pmatrix} \underset{\sim}{X}_1 \\ \underset{\sim}{X}_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \underset{\sim}{\mu}_1 \\ \underset{\sim}{\mu}_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right)$

Marginally : $\underset{\sim}{X}_i \sim N(\mu_i, \Sigma_{ii})$

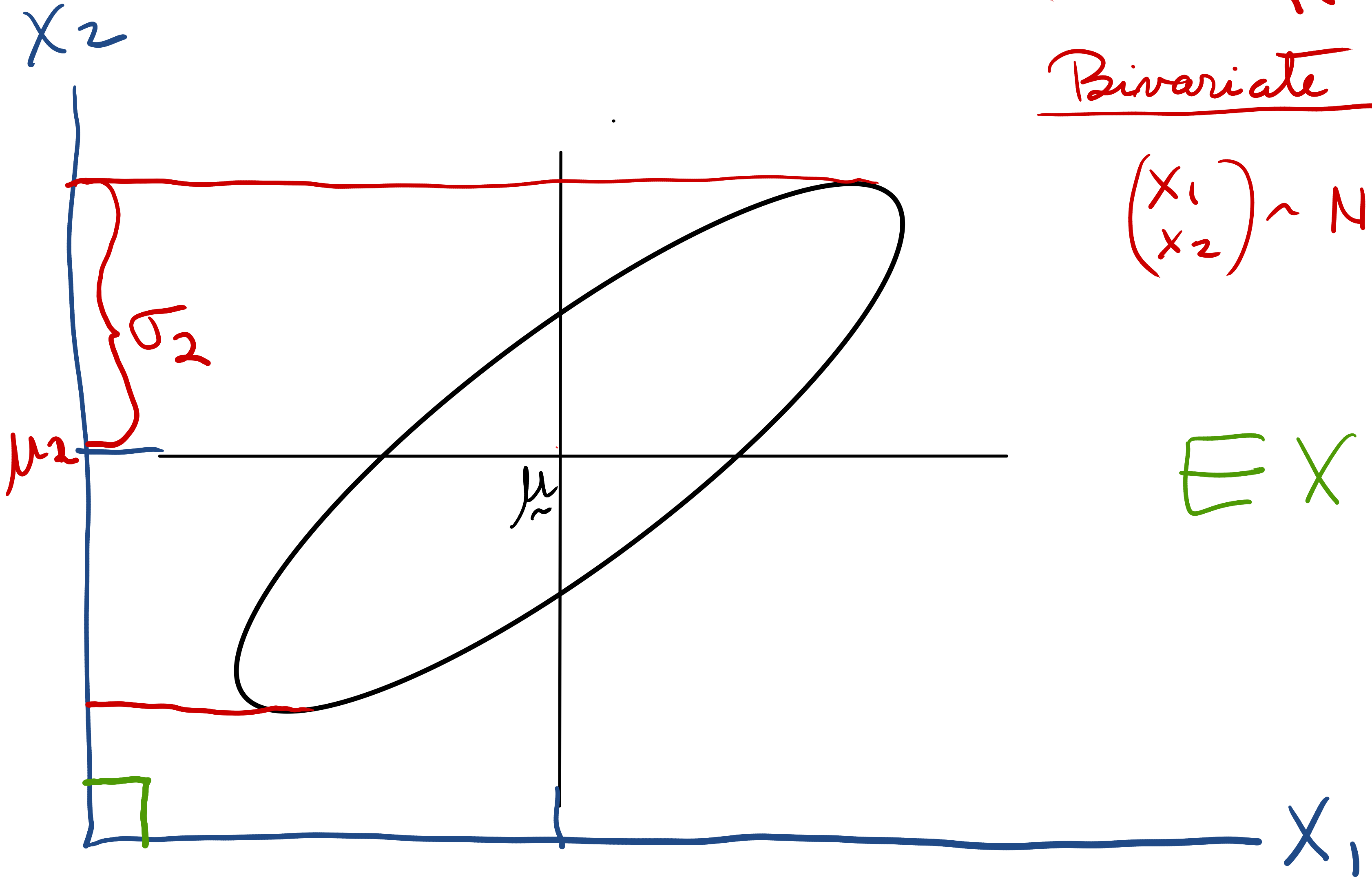
Conditionally $X_2 \mid X_1 = \underset{\sim}{x}_1 \sim N\left(\mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(x_1 - \underset{\sim}{\mu}_1),\ \underbrace{\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}}_{\text{Schur complement of } \Sigma_{22} \text{ in } \Sigma}\right)$

The Shape of Ellipses

$$\begin{pmatrix} \underset{\sim}{X}_1 \\ \underset{\sim}{X}_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right)$$

Bivariate

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}\right)$$



EX

Prove

The Shape of Ellipses

$$\begin{pmatrix}\tilde{X}_1\\ \tilde{X}_2\end{pmatrix} \sim N\left(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}, \begin{pmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{pmatrix}\right)$$

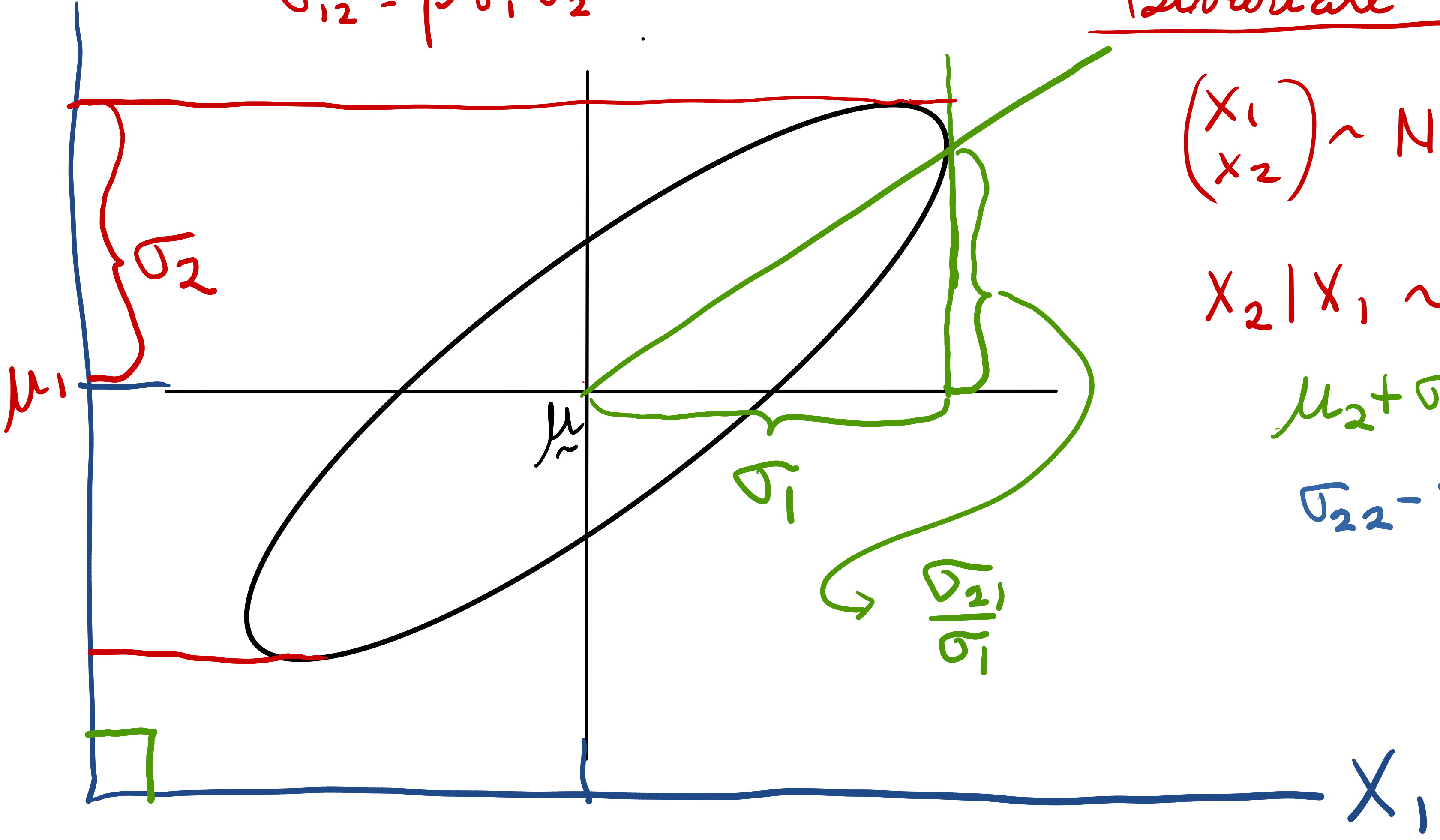
$\sigma_1 = \sqrt{\sigma_{11}}$ $\quad \sigma_2 = \sqrt{\sigma_{22}}$

$\sigma_{12} = \rho\,\sigma_1\sigma_2$

Bivariate

$$\begin{pmatrix}X_1\\ X_2\end{pmatrix} \sim N\left(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}, \begin{pmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{21} & \sigma_{22}\end{pmatrix}\right)$$

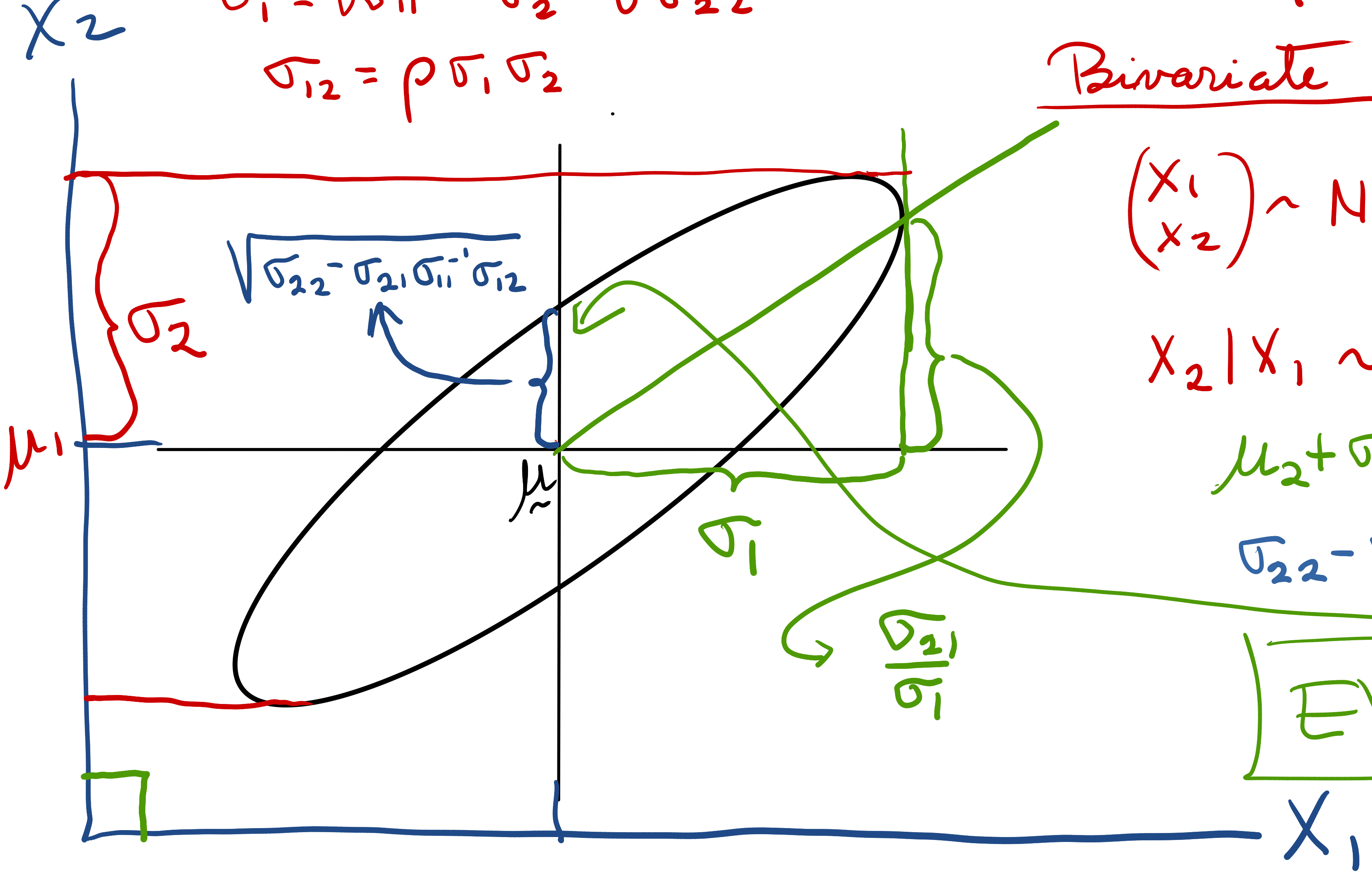
$$X_2 \mid X_1 \sim N\left(\mu_2 + \sigma_{21}\sigma_{11}^{-1}(x_1 - \mu_1),\ \sigma_{22} - \sigma_{21}\sigma_{11}^{-1}\sigma_{12}\right)$$



The Shape of Ellipses

$\sigma_1 = \sqrt{\sigma_{11}} \qquad \sigma_2 = \sqrt{\sigma_{22}}$

$\sigma_{12} = \rho\,\sigma_1\sigma_2$



$$\begin{pmatrix} \underset{\sim}{X}_1 \\ \underset{\sim}{X}_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right)$$

Bivariate

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}\right)$$

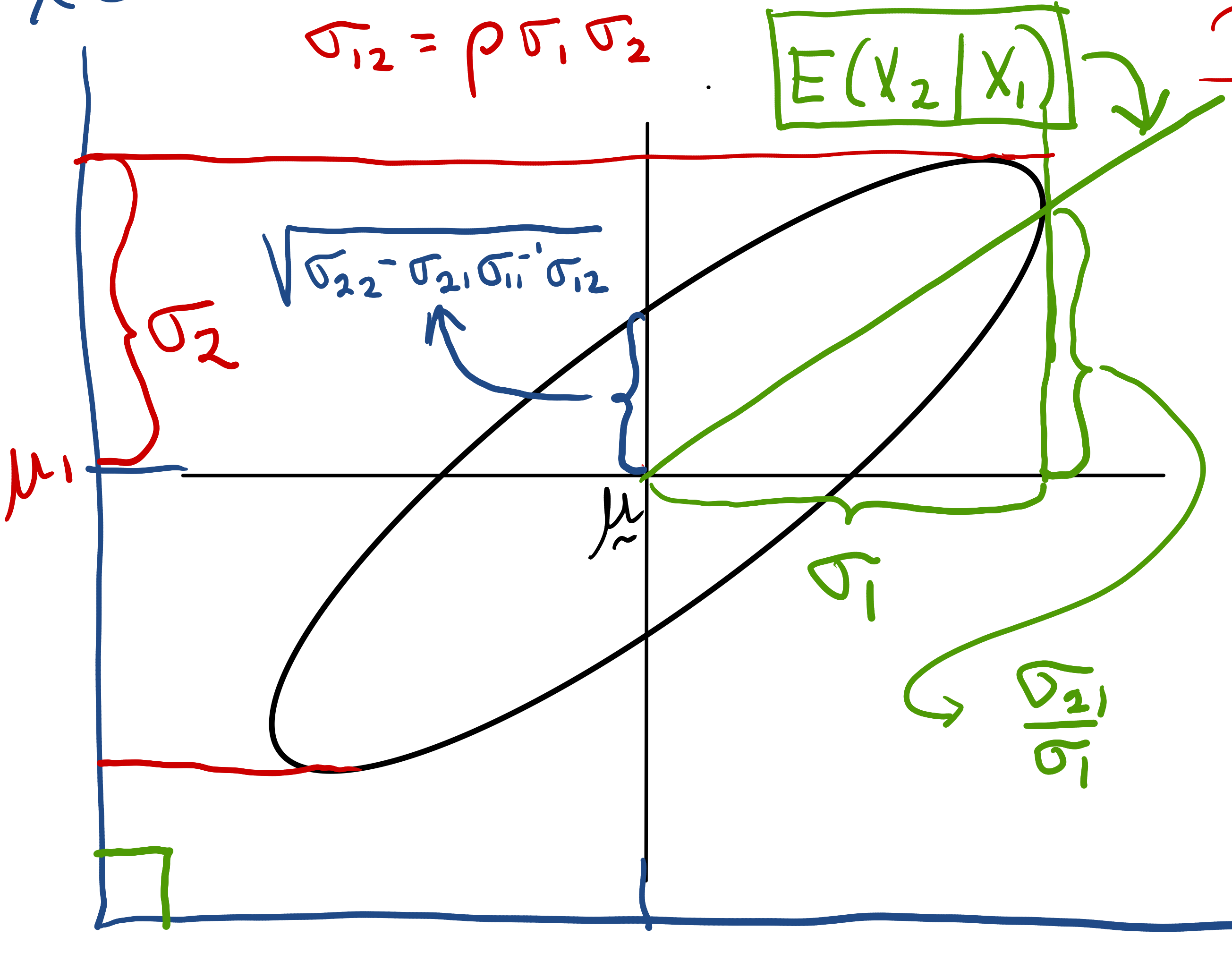
$$X_2 \mid X_1 \sim N\left(\mu_2 + \sigma_{21}\sigma_{11}^{-1}(x_1 - \mu_1),\; \sigma_{22} - \sigma_{21}\sigma_{11}^{-1}\sigma_{12}\right)$$

$\boxed{\text{EX 7: Prove}}$

The Shape of Ellipses

$\sigma_1 = \sqrt{\sigma_{11}} \qquad \sigma_2 = \sqrt{\sigma_{22}}$

X_2



X_1

$$\begin{pmatrix} \underset{\sim}{X}_1 \\ \underset{\sim}{X}_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \right)$$

Bivariate

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \right)$$

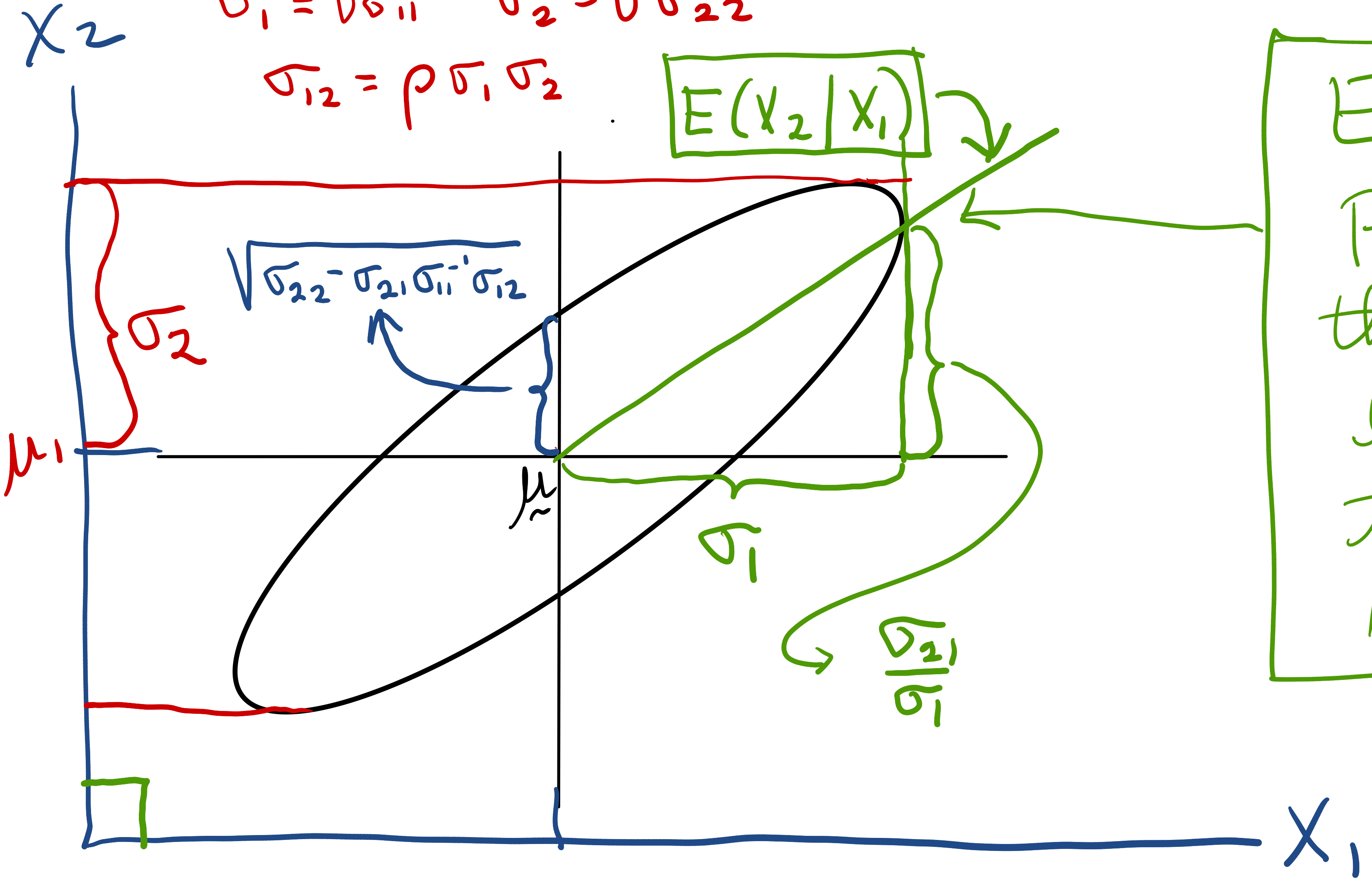
$$X_2 \mid X_1 \sim N\Big(\mu_2 + \sigma_{21}\sigma_{11}^{-1}(x_1 - \mu_1),\ \sigma_{22} - \sigma_{21}\sigma_{11}^{-1}\sigma_{12} \Big)$$

- Marginal SD = shadow of E
- conditional SD = slice of E in center

The Shape of Ellipses

$\sigma_1 = \sqrt{\sigma_{11}} \quad \sigma_2 = \sqrt{\sigma_{22}}$

$\sigma_{12} = \rho\,\sigma_1\sigma_2$



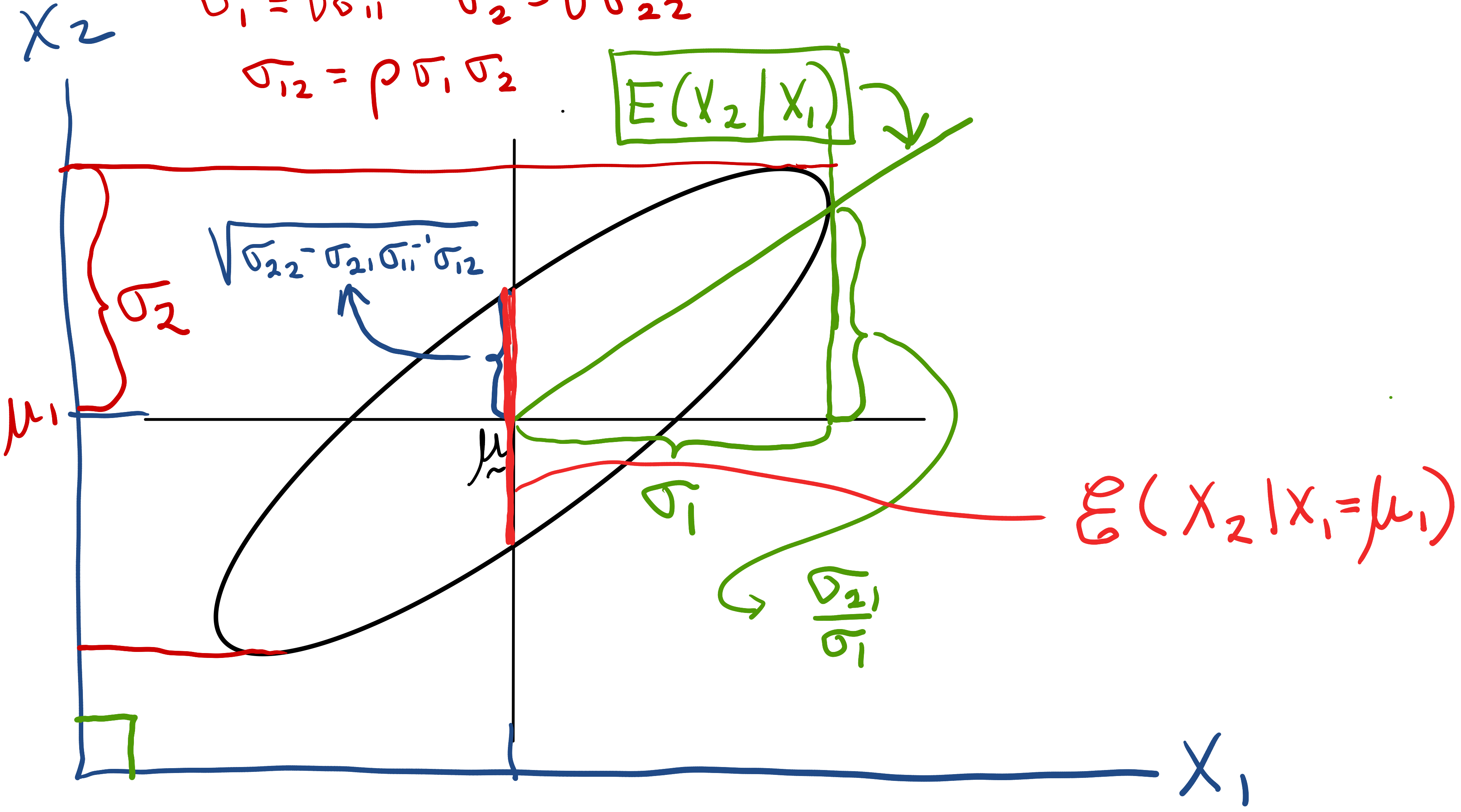
EX 8

Prove that the regression line goes through the point of vertical tangency

The Shape of Ellipses

$\sigma_1 = \sqrt{\sigma_{11}}$ $\sigma_2 = \sqrt{\sigma_{22}}$

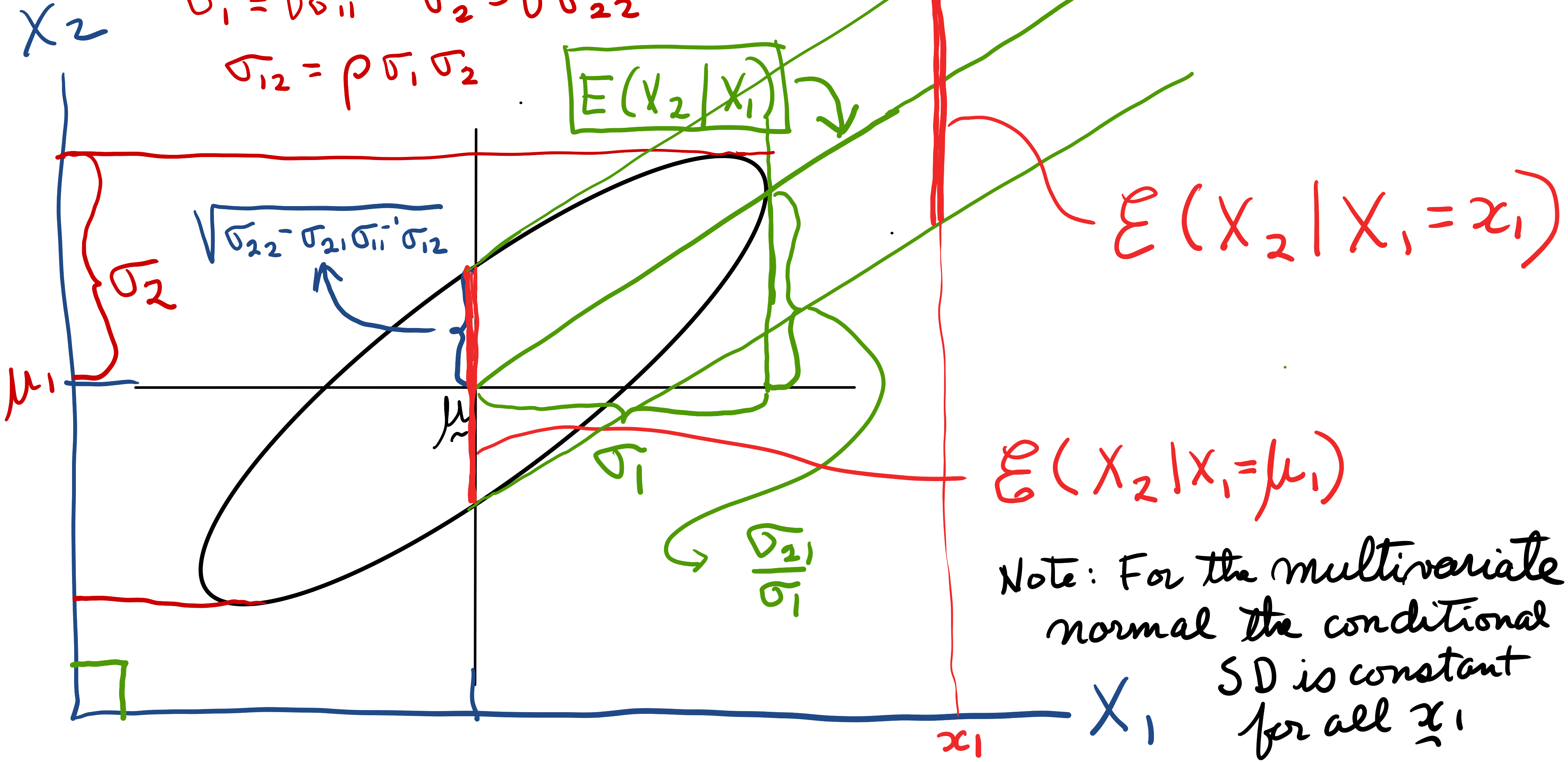
$\sigma_{12} = \rho\sigma_1\sigma_2$



The Shape of Ellipses

$\sigma_1 = \sqrt{\sigma_{11}}$ $\quad \sigma_2 = \sqrt{\sigma_{22}}$

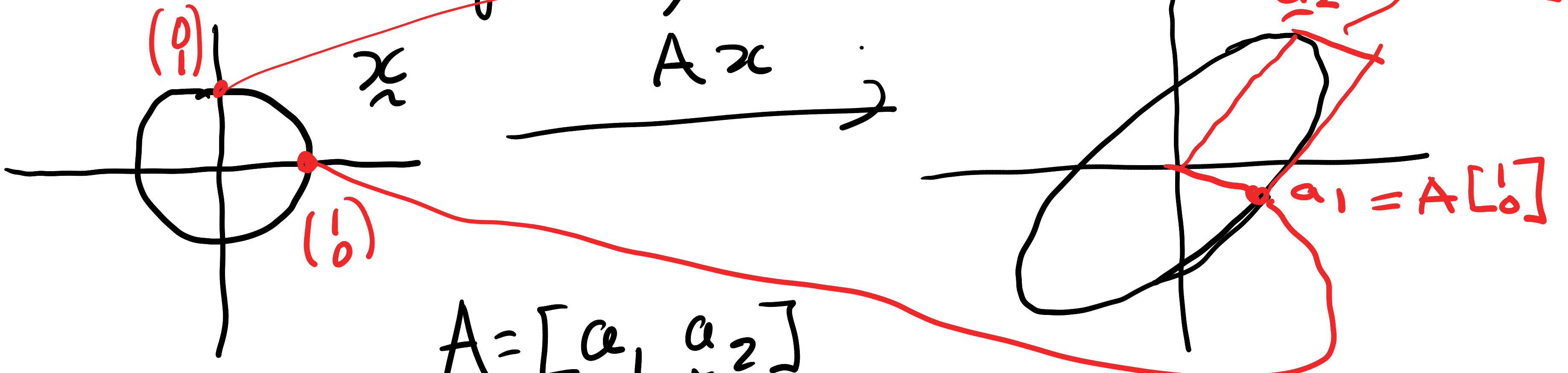
$\sigma_{12} = \rho\,\sigma_1\sigma_2$



Note: For the multivariate normal the conditional SD is constant for all $\underset{\sim}{x}_1$

Drawing ellipses

- All ellipses are affine transformations of a unit circle (sphere)



$$A = [\underset{\sim}{a}_1, \underset{\sim}{a}_2]$$

More generally $\mathcal{E} = \underset{\sim}{c} + A\,\mathcal{U}$

center　　unit sphere

EX 9
Write an R program to draw ellipses

Extended Cauchy-Schwarz Theorem

Let $A_{n\times n}$ be PD (positive definite $\Rightarrow$ symmetric)

Then $\forall\ \underset{\sim}{x}, \underset{\sim}{y} \in \mathbb{R}^n$

$$(\underset{\sim}{x}'\underset{\sim}{y})^2 \leq (\underset{\sim}{x}' A \underset{\sim}{x})(\underset{\sim}{y}' A^{-1} \underset{\sim}{y})$$

with $=$ iff $\exists\ a, b$, not both 0

$\ni\ aAx + by = 0$

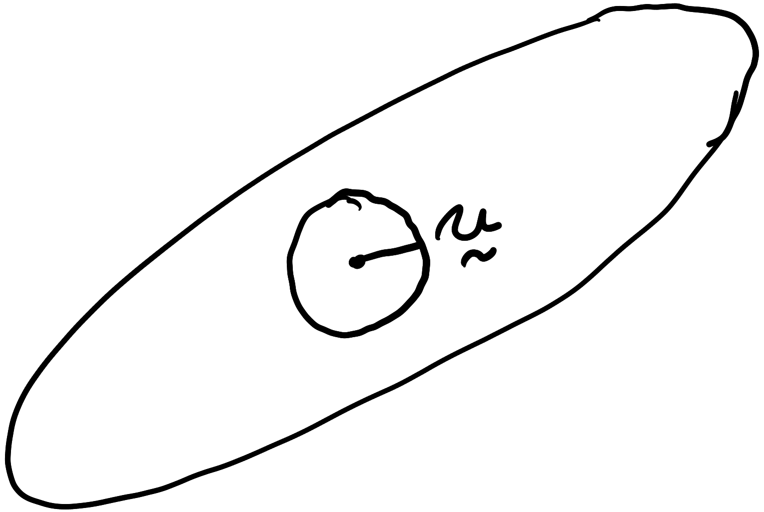
equivalent conditions
$\left\{\begin{array}{l} \text{iff } \ni\ ax + bA^{-1}y = 0\ ) \\ \text{iff } \underset{\sim}{0}, A\underset{\sim}{x}, y \text{ are collinear} \\ \text{iff } \underset{\sim}{0}, \underset{\sim}{x}, A^{-1}\underset{\sim}{y} \text{ '' \quad ''} \end{array}\right.$

$\boxed{\text{Ex 10 Prove}}$

Shadow / Slice Theorem

Let $\mathcal{E} = \{ \underset{\sim}{x} : \underset{\sim}{x}' A^{-1} \underset{\sim}{x} = 1 \}$ with A PD.

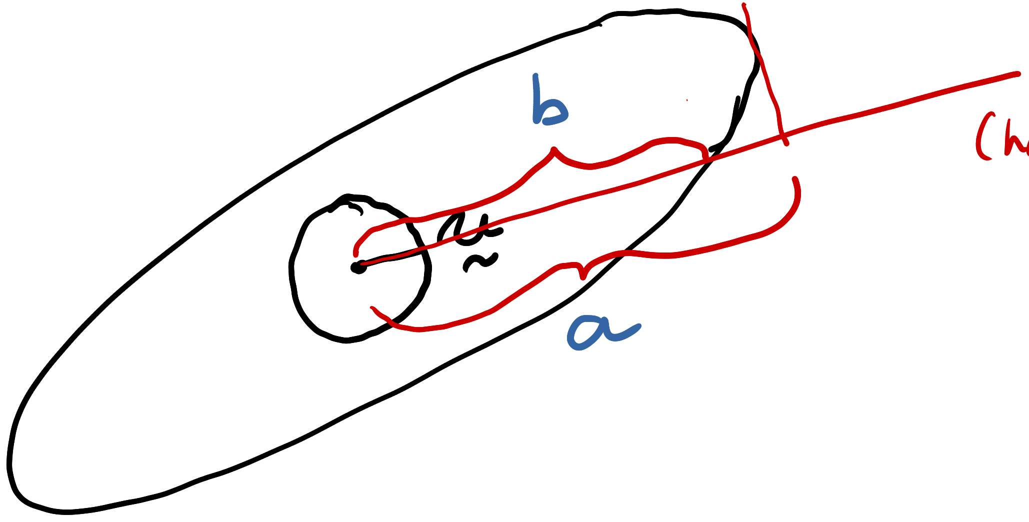
Let $\underset{\sim}{u}$ be a unit vector



Shadow / Slice Theorem

Let $\mathcal{E} = \{ \underset{\sim}{x} : \underset{\sim}{x}' A^{-1} \underset{\sim}{x} = r^2 \}$ with A PD.

Let $\underset{\sim}{u}$ be a unit vector



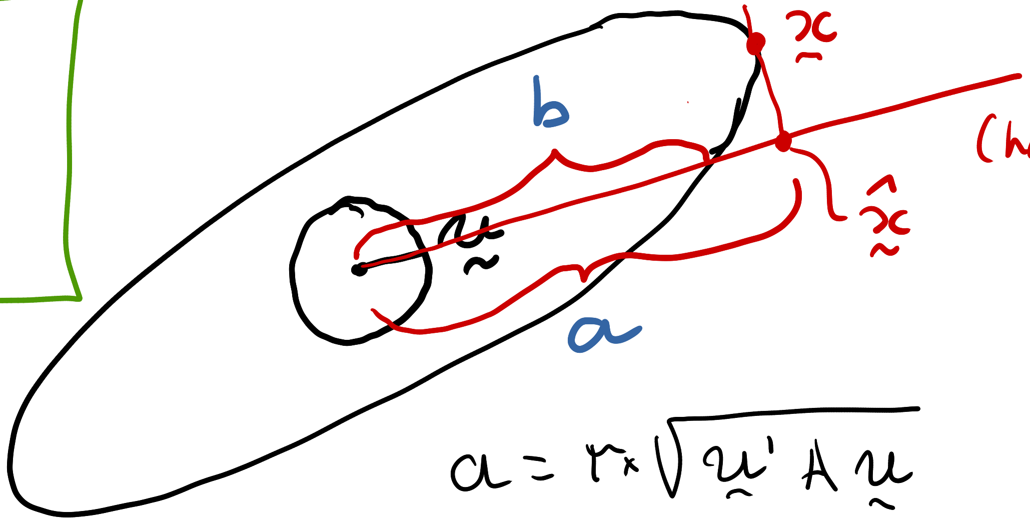
Let a be length of the perpendicular (half-) shadow of $\mathcal{E}$ on span($\underset{\sim}{u}$) and let b be the length of the half-"slice" of $\mathcal{E}$ along span ($\underset{\sim}{u}$)

Shadow / Slice Theorem

Let $\mathcal{E} = \{ \underset{\sim}{x} : \underset{\sim}{x}' A^{-1} \underset{\sim}{x} = r^2 \}$ with A PD.

Let $\underset{\sim}{u}$ be a unit vector

Let a be length of the perpendicular (half-) shadow of $\mathcal{E}$ on span($\underset{\sim}{u}$) and let b be the length of the half-"slice" of $\mathcal{E}$ along span($\underset{\sim}{u}$)



EX
Prove

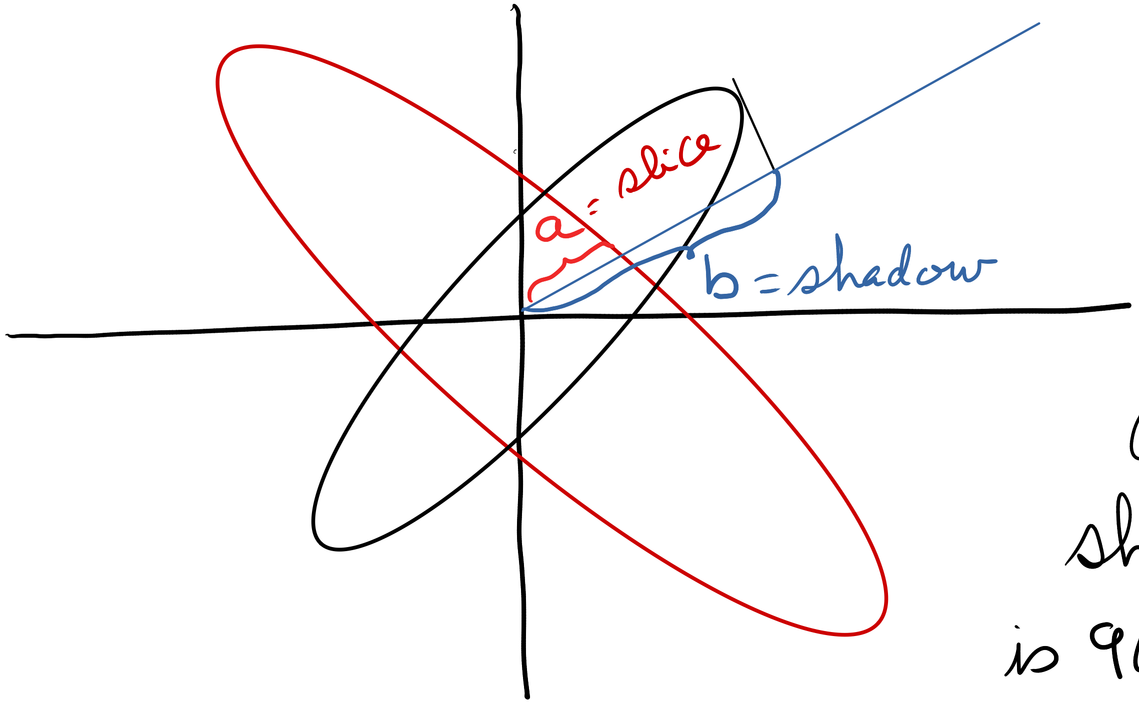
$$a = r \times \sqrt{\underset{\sim}{u}' A \underset{\sim}{u}}$$

$$b = r \times \frac{1}{\sqrt{\underset{\sim}{u}' A^{-1} \underset{\sim}{u}}}$$

Relationship between

$$\mathcal{E} = \{ \underset{\sim}{x} : \underset{\sim}{x}' A^{-1} \underset{\sim}{x} = 1 \}$$

and $\mathcal{D} = \{ \underset{\sim}{y} : \underset{\sim}{y}' A \, \underset{\sim}{y} = 1 \}$



$$a \, b = 1$$

On $\mathbb{R}^2$
shape of $\mathcal{D}$
is $90°$ rotation of
shape of $\mathcal{E}$

Relationship between

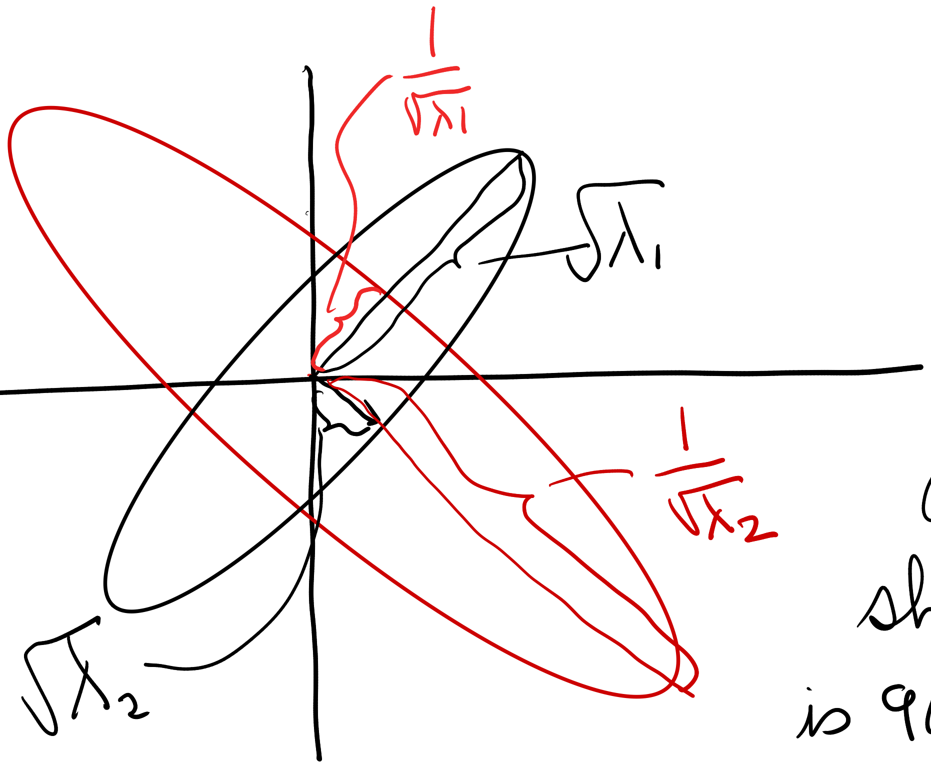
$$\mathcal{E} = \{\underset{\sim}{x} : \underset{\sim}{x}' A^{-1} \underset{\sim}{x} = 1\}$$

and $$\mathcal{D} = \{\underset{\sim}{y} : \underset{\sim}{y}' A \underset{\sim}{y} = 1\}$$

If $A = \Gamma \Lambda \Gamma'$

$$= [\underset{\sim}{\gamma}_1 \ \underset{\sim}{\gamma}_2] \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \Gamma'$$

$$\lambda_1 \ge \lambda_2 > 0$$



$ab = 1$

On $\mathbb{R}^2$ shape of $\mathcal{D}$ is $90°$ rotation of shape of $\mathcal{E}$

GLH : General Linear Hypothesis

Linear model $\underset{n\times 1}{\tilde{Y}} = \underset{n\times p}{X}\underset{p\times 1}{\tilde{\beta}} + \tilde{\varepsilon}$ $\quad \tilde{\varepsilon} \sim N(\tilde{0}, \sigma^2 I)$

Test: $H_0 : \tilde{\eta} = \underset{h\times p}{L}\tilde{\beta} = \tilde{0}$

Use $\hat{\eta} = L\hat{\beta}$

$$Var(\hat{\eta}) = L\,Var(\hat{\beta})\,L'$$
$$= \sigma^2 L(X'X)^{-1}L'$$

$$SSH = \hat{\eta}'\left(Var(\hat{\eta})\right)^{-1}\hat{\eta} = (L\hat{\beta})'(L(X'X)^{-1}L')^{-1}L\hat{\beta}$$

If L is of full row rank & H_0 true $\dfrac{SSH/h}{SSE/\nu} \sim F_{h,\nu}$

$\nu = n - p$

$\boxed{\text{EX 11}}$ Show that a hypothesis depends only on the row space of L

i.e. if L_1 and L_2 have full row rank and have the same row space, then

$$SSH(L_1) = SSH(L_2)$$

Woodbury Formula

Let $\underset{n\times n}{A}$ and $\underset{k\times k}{C}$ be matrices. then

$$(A+UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1}+VA^{-1}U)^{-1}VA^{-1}$$

if the inverses exist.

Key Theorem to remember

Marg Var = Mean Condl Var + Var Condl Mean

Condl = "Conditional"

Sorry!

$$\cancel{E(Y)}\ \mathrm{Var}(Y) = E(\mathrm{Var}(Y \mid X)) + \mathrm{Var}(E(Y \mid X))$$

$\boxed{\text{EX 12 Prove}}$

$$E(Y) = E(E(Y \mid X))$$