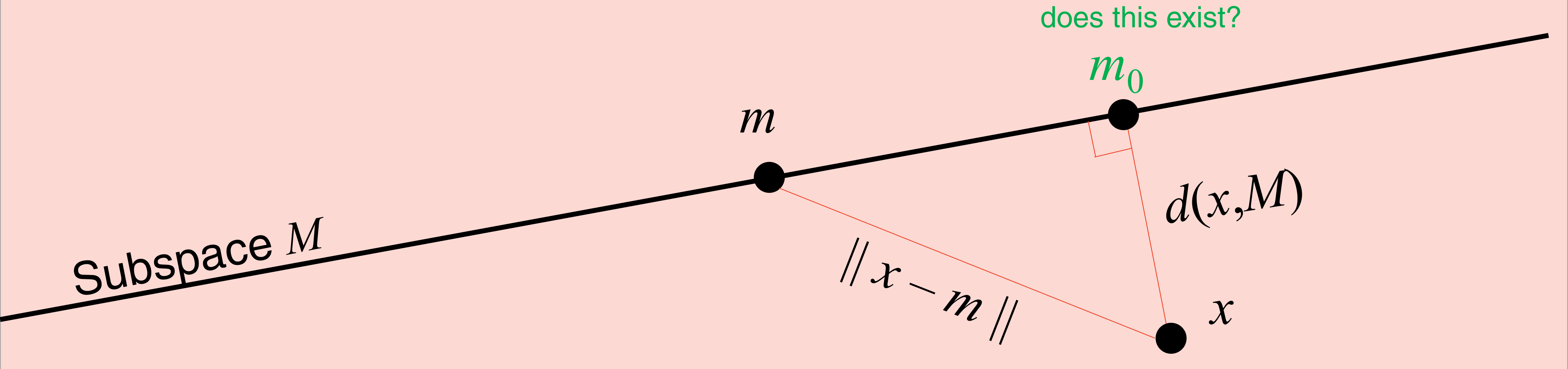
Mathematics for Robotics (ROB-GY 6013 Section A)

- **Week 9:**
  - Theme: Least Squares
    - Using the Projection Theorem: Normal Equations
    - Quadratic Forms

Recap: Without Classical Projection Theorem

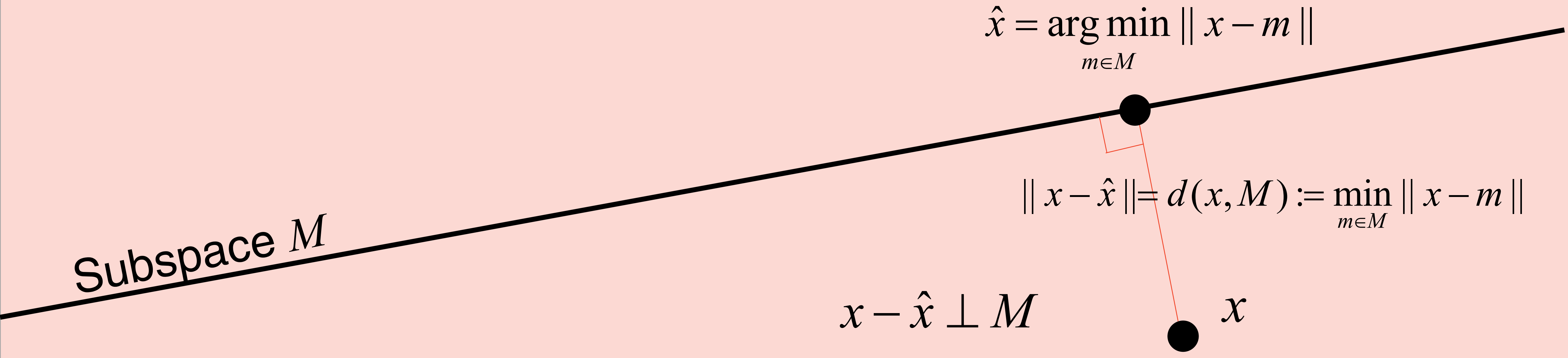
Let $(\mathcal{X}, \mathcal{F}, \langle \cdot , \cdot \rangle)$ be a **inner product space**, M be a **subspace** of $\mathcal{X}$, and x be an arbitrary **point** in $\mathcal{X}$.



Recap: Classical Projection Theorem

Let $(\mathcal{X}, \mathcal{F}, \langle \cdot , \cdot \rangle)$ be a **finite-dimensional** (real) **inner product space**, M be a **subspace** of $\mathcal{X}$, and x be an arbitrary **point** in $\mathcal{X}$.

Inner product space $(\mathcal{X}, \mathcal{F}, \langle \cdot , \cdot \rangle)$



Recap: Classical Projection Theorem

- Let $(\mathcal{X}, \mathcal{F}, \langle \cdot , \cdot \rangle)$ be a finite-dimensional (real) **inner product space** and M a **subspace** of $\mathcal{X}$. Then, $\forall x \in \mathcal{X}$, $\exists$ **unique** $\hat{x} \in M$ such that

$$\|x-\hat{x}\| = d(x,M) := \inf_{m\in M} \|x-m\| = \min_{m\in M} \|x-m\|$$

  where we can write **min**imum instead of **inf**imum, because the infimum is achieved. Moreover $\hat{x} \in M$ is characterized by $x-\hat{x} \perp M$

- Use of $\hat{x}$ (best approximation) instead of m_0 because this is the standard notation for estimation problems.

Find $\hat{x}$ with Projection Theorem

- By Projection Theorem, $\hat{x}$ **exists** and is **unique** (characterized by $x - \hat{x} \perp M$)

Find $\hat{x}$ with Projection Theorem

- By Projection Theorem, $\hat{x}$ **exists** and is **unique** (characterized by $x - \hat{x} \perp M$)
- Let $M = \text{span}\{y^1, \dots, y^k\}$: $\hat{x} = \alpha_1 y^1 + \alpha_2 y^2 + \dots + \alpha_k y^k$

Find $\hat{x}$ with Projection Theorem

- By Projection Theorem, $\hat{x}$ **exists** and is **unique** (characterized by $x - \hat{x} \perp M$)
- Let $M = \text{span}\{y^1, \ldots, y^k\}$: $\hat{x} = \alpha_1 y^1 + \alpha_2 y^2 + \ldots + \alpha_k y^k$
- Then: $(x - \hat{x} \perp M)$

$$\Updownarrow$$

$$x - \hat{x} \perp y^i \quad 1 \leq i \leq k$$

Find $\hat{x}$ with Projection Theorem

- By Projection Theorem, $\hat{x}$ **exists** and is **unique** (characterized by $x-\hat{x}\perp M$)
- Let $M = \mathrm{span}\{y^1, \dots, y^k\}$: $\hat{x}=\alpha_1 y^1+\alpha_2 y^2+\ldots+\alpha_k y^k$
- Then: $(x-\hat{x}\perp M)$

$$\Updownarrow$$

$$x-\hat{x}\perp y^i \quad 1\le i\le k$$

$$\Updownarrow$$

$$\left\langle x-\hat{x}, y^i\right\rangle=0 \quad 1\le i\le k$$

Find $\hat{x}$ with Projection Theorem

- By Projection Theorem, $\hat{x}$ **exists** and is **unique** (characterized by $x - \hat{x} \perp M$)
- Let $M = \operatorname{span}\{y^1, \ldots, y^k\}$: $\hat{x} = \alpha_1 y^1 + \alpha_2 y^2 + \ldots + \alpha_k y^k$
- Then: $(x - \hat{x} \perp M)$

$$\Updownarrow$$

$$x - \hat{x} \perp y^i \quad 1 \le i \le k$$

$$\Updownarrow$$

$$\left\langle x - \hat{x}, y^i \right\rangle = 0 \quad 1 \le i \le k \qquad\qquad \left\langle x - \hat{x}, y^i \right\rangle = \left\langle \hat{x}, y^i \right\rangle - \left\langle x, y^i \right\rangle$$

$$\Updownarrow$$

$$\left\langle \hat{x}, y^i \right\rangle = \left\langle x, y^i \right\rangle \quad 1 \le i \le k$$

Find $\hat{x}$ with Projection Theorem

- By Projection Theorem, $\hat{x}$ **exists** and is **unique** (characterized by $x - \hat{x} \perp M$)
- Let $M = \text{span}\{y^1, \dots, y^k\}$: $\hat{x} = \alpha_1 y^1 + \alpha_2 y^2 + \ldots + \alpha_k y^k$
- Then: $(x - \hat{x} \perp M)$

$$\Updownarrow$$

$$x - \hat{x} \perp y^i \quad 1 \le i \le k$$

$$\Updownarrow$$

$$\left\langle x - \hat{x}, y^i \right\rangle = 0 \quad 1 \le i \le k \qquad\qquad \left\langle x - \hat{x}, y^i \right\rangle = \left\langle \hat{x}, y^i \right\rangle - \left\langle x, y^i \right\rangle$$

$$\Updownarrow$$

$$\boxed{\left\langle \hat{x}, y^i \right\rangle = \left\langle x, y^i \right\rangle \quad 1 \le i \le k}$$

$$
\begin{gathered}
\langle \widehat{x}, y^i \rangle = \langle x, y^i \rangle \; 1 \leq i \leq k \\
\Updownarrow \\
\langle \alpha_1 y^1 + \alpha_2 y^2 + \cdots + \alpha_k y^k, y^i \rangle = \langle x, y^i \rangle \; 1 \leq i \leq k \\
\Updownarrow \\
\alpha_1 \langle y^1, y^i \rangle + \alpha_2 \langle y^2, y^i \rangle + \cdots + \alpha_k \langle y^k, y^i \rangle = \langle x, y^i \rangle \; 1 \leq i \leq k
\end{gathered}
$$

- Find α's

$$\langle \widehat{x}, y^i \rangle = \langle x, y^i \rangle \; 1 \le i \le k$$
$$\Updownarrow$$
$$\langle \alpha_1 y^1 + \alpha_2 y^2 + \cdots + \alpha_k y^k, y^i \rangle = \langle x, y^i \rangle \; 1 \le i \le k$$
$$\Updownarrow$$
$$\alpha_1 \langle y^1, y^i \rangle + \alpha_2 \langle y^2, y^i \rangle + \cdots + \alpha_k \langle y^k, y^i \rangle = \langle x, y^i \rangle \; 1 \le i \le k$$

- Find α's

$$\langle \widehat{x}, y^i \rangle = \langle x, y^i \rangle \; 1 \le i \le k$$
$$\Updownarrow$$
$$\langle \alpha_1 y^1 + \alpha_2 y^2 + \cdots + \alpha_k y^k, y^i \rangle = \langle x, y^i \rangle \; 1 \le i \le k$$
$$\Updownarrow$$
$$\alpha_1 \langle y^1, y^i \rangle + \alpha_2 \langle y^2, y^i \rangle + \cdots + \alpha_k \langle y^k, y^i \rangle = \langle x, y^i \rangle \; 1 \le i \le k$$

$$\boxed{\begin{array}{ll} i=1 & \alpha_1 \langle y^1, y^1 \rangle + \alpha_2 \langle y^2, y^1 \rangle + \cdots + \alpha_k \langle y^k, y^1 \rangle = \langle x, y^1 \rangle \\ i=2 & \alpha_1 \langle y^1, y^2 \rangle + \alpha_2 \langle y^2, y^2 \rangle + \cdots + \alpha_k \langle y^k, y^2 \rangle = \langle x, y^2 \rangle \\ \vdots & \vdots \\ i=k & \alpha_1 \langle y^1, y^k \rangle + \alpha_2 \langle y^2, y^k \rangle + \cdots + \alpha_k \langle y^k, y^k \rangle = \langle x, y^k \rangle \end{array}}$$

Normal Equations

$$G(y^1,\cdots,y^k) := \begin{bmatrix} \langle y^1,y^1\rangle & \langle y^1,y^2\rangle & \cdots & \langle y^1,y^k\rangle \\ \langle y^2,y^1\rangle & \langle y^2,y^2\rangle & \cdots & \langle y^2,y^k\rangle \\ \vdots & \vdots & & \vdots \\ \langle y^k,y^1\rangle & \langle y^k,y^2\rangle & \cdots & \langle y^k,y^k\rangle \end{bmatrix} \quad \alpha := \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{bmatrix}, \ \beta := \begin{bmatrix} \langle x,y^1\rangle \\ \langle x,y^2\rangle \\ \vdots \\ \langle x,y^k\rangle \end{bmatrix} =: \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix}$$

Gram Matrix

$G_{ij} := \langle y^i, y^j \rangle$

$$\boxed{\begin{array}{ll} i=1 & \alpha_1\langle y^1,y^1\rangle + \alpha_2\langle y^2,y^1\rangle + \cdots + \alpha_k\langle y^k,y^1\rangle = \langle x,y^1\rangle \\ i=2 & \alpha_1\langle y^1,y^2\rangle + \alpha_2\langle y^2,y^2\rangle + \cdots + \alpha_k\langle y^k,y^2\rangle = \langle x,y^2\rangle \\ \vdots & \qquad\qquad\qquad\qquad \vdots \\ i=k & \alpha_1\langle y^1,y^k\rangle + \alpha_2\langle y^2,y^k\rangle + \cdots + \alpha_k\langle y^k,y^k\rangle = \langle x,y^k\rangle \end{array}}$$

Normal Equations $\quad G^T\alpha = \beta$

$$G(y^1, \cdots, y^k) := \begin{bmatrix} \langle y^1, y^1 \rangle & \langle y^1, y^2 \rangle & \cdots & \langle y^1, y^k \rangle \\ \langle y^2, y^1 \rangle & \langle y^2, y^2 \rangle & \cdots & \langle y^2, y^k \rangle \\ \vdots & \vdots & & \vdots \\ \langle y^k, y^1 \rangle & \langle y^k, y^2 \rangle & \cdots & \langle y^k, y^k \rangle \end{bmatrix} \quad \alpha := \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{bmatrix}, \ \beta := \begin{bmatrix} \langle x, y^1 \rangle \\ \langle x, y^2 \rangle \\ \vdots \\ \langle x, y^k \rangle \end{bmatrix} =: \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix}$$

Gram Matrix

$G_{ij} := \langle y^i, y^j \rangle$

If real: $G^T = G$

$$\boxed{\begin{array}{ll} i = 1 & \alpha_1 \langle y^1, y^1 \rangle + \alpha_2 \langle y^2, y^1 \rangle + \cdots + \alpha_k \langle y^k, y^1 \rangle = \langle x, y^1 \rangle \\ i = 2 & \alpha_1 \langle y^1, y^2 \rangle + \alpha_2 \langle y^2, y^2 \rangle + \cdots + \alpha_k \langle y^k, y^2 \rangle = \langle x, y^2 \rangle \\ \vdots & \vdots \\ i = k & \alpha_1 \langle y^1, y^k \rangle + \alpha_2 \langle y^2, y^k \rangle + \cdots + \alpha_k \langle y^k, y^k \rangle = \langle x, y^k \rangle \end{array}}$$

Normal Equations $G^T \alpha = \beta$

$$G(y^1, \cdots, y^k) := \begin{bmatrix} \langle y^1, y^1 \rangle & \langle y^1, y^2 \rangle & \cdots & \langle y^1, y^k \rangle \\ \langle y^2, y^1 \rangle & \langle y^2, y^2 \rangle & \cdots & \langle y^2, y^k \rangle \\ \vdots & & & \vdots \\ \langle y^k, y^1 \rangle & \langle y^k, y^2 \rangle & \cdots & \langle y^k, y^k \rangle \end{bmatrix} \quad \alpha := \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{bmatrix}, \ \beta := \begin{bmatrix} \langle x, y^1 \rangle \\ \langle x, y^2 \rangle \\ \vdots \\ \langle x, y^k \rangle \end{bmatrix} =: \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix}$$

Is this invertible?

Gram Matrix

$G_{ij} := \langle y^i, y^j \rangle$

If real: $G^T = G$

$$\boxed{\begin{array}{ll} i=1 & \alpha_1 \langle y^1, y^1 \rangle + \alpha_2 \langle y^2, y^1 \rangle + \cdots + \alpha_k \langle y^k, y^1 \rangle = \langle x, y^1 \rangle \\ i=2 & \alpha_1 \langle y^1, y^2 \rangle + \alpha_2 \langle y^2, y^2 \rangle + \cdots + \alpha_k \langle y^k, y^2 \rangle = \langle x, y^2 \rangle \\ \vdots & \qquad\qquad\qquad\qquad \vdots \\ i=k & \alpha_1 \langle y^1, y^k \rangle + \alpha_2 \langle y^2, y^k \rangle + \cdots + \alpha_k \langle y^k, y^k \rangle = \langle x, y^k \rangle \end{array}}$$

Normal Equations

$$G^T \alpha = \beta$$

$$G(y^1, \cdots, y^k) := \begin{bmatrix} \langle y^1, y^1\rangle & \langle y^1, y^2\rangle & \cdots & \langle y^1, y^k\rangle \\ \langle y^2, y^1\rangle & \langle y^2, y^2\rangle & \cdots & \langle y^2, y^k\rangle \\ \vdots & & & \vdots \\ \langle y^k, y^1\rangle & \langle y^k, y^2\rangle & \cdots & \langle y^k, y^k\rangle \end{bmatrix} \quad \alpha := \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{bmatrix}, \ \beta := \begin{bmatrix} \langle x, y^1\rangle \\ \langle x, y^2\rangle \\ \vdots \\ \langle x, y^k\rangle \end{bmatrix} =: \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix}$$

Is this invertible?

Gram Matrix

$G_{ij} := <y^i, y^j>$

If real: $G^T = G$

Normal Equations $$G^T \alpha = \beta$$

Proposition 3.41 (Invertibility of the Gram Matrix)

- Let $g(y^1, \dots, y^k) := \det G(y^1, \dots, y^k)$ be the **determinant** of the **Gram Matrix**.

  Then $g(y^1, \dots, y^k) \neq 0$ if and only if $\{y^1, \dots, y^k\}$ is **linearly independent**.

Normal Equations

- Seek $\hat{x} = \arg\min_{m \in M} \|x - m\|$
- Pick basis
  - Let $M = \text{span}\{y^1, \ldots, y^k\}$:

$$\hat{x} = \alpha_1 y^1 + \alpha_2 y^2 + \ldots + \alpha_k y^k$$

- Use normal equations

$$G^T \alpha = \beta$$

- Get α's

Linear Regression (Overdetermined Equations)

- Seek $\hat{x} = \arg\min_{m \in M} \|x - m\|$

- Pick basis
  - Let $M = \text{span}\{y^1, \dots, y^k\}$:

    $\hat{x} = \alpha_1 y^1 + \alpha_1 y^1 + \ldots + \alpha_k y^k$

- Use normal equations

  $$G^T \alpha = \beta$$

- Get α's

Linear Regression (Overdetermined Equations)

- Seek $\hat{x} = \underset{m \in M}{\arg\min} \|x - m\|$

- Pick basis
  - Let $M = \operatorname{span}\{y^1, \dots, y^k\}$:

    $\hat{x} = \alpha_1 y^1 + \alpha_1 y^1 + \ldots + \alpha_k y^k$

- Use normal equations

  $$G^T \alpha = \beta$$

- Get α's

Linear Regression (Overdetermined Equations)

- Seek $\hat{x} = \arg\min_{m \in M} \|x - m\|$
- Pick basis
  - Let $M = \text{span}\{y^1, \dots, y^k\}$:

    $\hat{x} = \alpha_1 y^1 + \alpha_1 y^1 + \ldots + \alpha_k y^k$
- Use normal equations

  $$G^T \alpha = \beta$$
- Get α's

Wealth (USD)	Sleep (Avg Hrs/day)	Happiness (0, 100)
1	1	0
100	2	10
25	5	50
50	8	100

Linear Regression (Overdetermined Equations)

- Seek $\hat{x} = \arg\min_{m \in M} \|x - m\|$

- Pick basis
  - Let $M = \text{span}\{y^1, \ldots, y^k\}$:

    $\hat{x} = \alpha_1 y^1 + \alpha_1 y^1 + \ldots + \alpha_k y^k$

- Use normal equations

  $$G^T \alpha = \beta$$

- Get α's

Wealth (USD)	Sleep (Avg Hrs/day)	Happiness (0, 100)
1	1	0
100	2	10
25	5	50
50	8	100

$$A\hat{\alpha} \approx b$$

$$\begin{bmatrix} 1 & 1 \\ 100 & 2 \\ 25 & 5 \\ 50 & 8 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \approx \begin{bmatrix} 0 \\ 10 \\ 50 \\ 100 \end{bmatrix}$$

Linear Regression (Overdetermined Equations)

- **Seek** $\hat{x} = \arg\min_{m \in M} \|x - m\| = \arg\min_{A\hat{\alpha} \in M} \|b - A\hat{\alpha}\|$

- Pick basis
  - Let $M = \operatorname{span}\{y^1, \ldots, y^k\}$:

    $\hat{x} = \alpha_1 y^1 + \alpha_1 y^1 + \ldots + \alpha_k y^k$

- Use normal equations

  $$G^T \alpha = \beta$$

- Get α's

Linear Regression (Overdetermined Equations)

- **Seek** $\hat{x} = \arg\min_{m \in M} \|x - m\| = \arg\min_{A\hat{\alpha} \in M} \|b - A\hat{\alpha}\|$
- **Pick basis (columns of A)**
  - Let $M = \operatorname{span}\{y^1, \ldots, y^k\}$:

    $\hat{x} = \alpha_1 y^1 + \alpha_1 y^1 + \ldots + \alpha_k y^k$
- Use normal equations

  $$G^T \alpha = \beta$$
- Get α's

Linear Regression (Overdetermined Equations)

- **Seek** $\hat{x} = \arg\min_{m \in M} \|x - m\| = \arg\min_{A\hat{\alpha} \in M} \|b - A\hat{\alpha}\|$

- **Pick basis (columns of A)**
  - Let $M = \operatorname{span}\{y^1, \ldots, y^k\}$:

    $\hat{x} = \alpha_1 y^1 + \alpha_1 y^1 + \ldots + \alpha_k y^k$

- **Use normal equations**

$$G^T \alpha = \beta$$

- Get α's

Linear Regression (Overdetermined Equations)

- **Seek** $\hat{x} = \arg\min_{m \in M} \|x - m\| = \arg\min_{A\hat{\alpha} \in M} \|b - A\hat{\alpha}\|$
- **Pick basis (columns of A)**
  - Let $M = \mathrm{span}\{y^1, \ldots, y^k\}$:

    $\hat{x} = \alpha_1 y^1 + \alpha_1 y^1 + \ldots + \alpha_k y^k$
- **Use normal equations**

  $$G^T \alpha = \beta$$
- **Get α's**

Remarks

- No vector calculus!

$$\widehat{\alpha} = \arg\min_{\alpha \in \mathbb{R}^m} \sqrt{(A\alpha - b)^\top (A\alpha - b)} = \arg\min_{\alpha \in \mathbb{R}^m} (A\alpha - b)^\top (A\alpha - b)$$

$$\frac{\partial}{\partial \alpha} \left((A\alpha - b)^\top (A\alpha - b)\right) = A^\top \cdot (A\alpha - b)$$

$$A^\top \cdot (A\widehat{\alpha} - b) = 0 \iff (A^\top A)\widehat{\alpha} = A^\top b \iff \widehat{\alpha} = (A^\top A)^{-1} A^\top b$$

Remarks

- No vector calculus!

- We

Remarks

- No vector calculus!

- We

Definition: Orthogonal Projection Operator

- Let $(\mathcal{X}, \mathcal{F}, \langle \cdot , \cdot \rangle)$ be a finite-dimensional (real) **inner product space** and M be a **subspace** of $\mathcal{X}$. For $x \in \mathcal{X}$ and $\hat{x} \in M$, the Projection Theorem shows the TFAE:

$$(a)\ \ x - \widehat{x} \perp M.$$

$$(b)\ \ \exists m^{\perp} \in M^{\perp} \textit{ such that } x = \widehat{x} + m^{\perp}.$$

$$(c)\ \ \|x - \widehat{x}\| = d(x, M) = \inf_{m \in M} \|x - m\|.$$

Definition: Orthogonal Projection Operator

- Let $(\mathcal{X}, \mathcal{F}, \langle \cdot , \cdot \rangle)$ be a finite-dimensional (real) **inner product space** and M be a **subspace** of $\mathcal{X}$. For $x \in \mathcal{X}$ and $\hat{x} \in M$, the Projection Theorem shows the TFAE:

$$(a)\ \ x - \widehat{x} \perp M.$$

$$(b)\ \ \exists m^{\perp} \in M^{\perp} \textit{ such that } x = \widehat{x} + m^{\perp}.$$

$$(c)\ \ \|x - \widehat{x}\| = d(x, M) = \inf_{m \in M} \|x - m\|.$$

- A function $P\colon \mathcal{X} \to M$ by $P(x) = \hat{x}$, where $\hat{x}$ satisfies any one of (a), (b), or (c), is called the **orthogonal projection** of $\mathcal{X}$ onto M.

Definition: Orthogonal Projection Operator

- Let $(\mathcal{X}, \mathcal{F}, \langle \cdot , \cdot \rangle)$ be a finite-dimensional (real) **inner product space** and M be a **subspace** of $\mathcal{X}$. For $x \in \mathcal{X}$ and $\hat{x} \in M$, the Projection Theorem shows the TFAE:

$$(a)\ \ x - \widehat{x} \perp M.$$

$$(b)\ \ \exists m^{\perp} \in M^{\perp}\ \textit{such that}\ x = \widehat{x} + m^{\perp}.$$

$$(c)\ \ \|x - \widehat{x}\| = d(x, M) = \inf_{m \in M} \|x - m\|.$$

- A function $P: \mathcal{X} \to M$ by $P(x) = \hat{x}$, where $\hat{x}$ satisfies any one of (a), (b), or (c), is called the **orthogonal projection** of $\mathcal{X}$ onto M.
  - P is a linear operator
  - Let $\{v^1, \ldots, v^k\}$ be an **orthonormal basis** for M. Then $P(x) = \sum_{i=1}^{k} \langle x, v^i \rangle v^i.$

Many Claims about Symmetric and Orthogonal Matrices

If A is $n \times n$ and real, then its eigenvalues and eigenvectors occur in complex conjugate pairs.

If A is $n \times n$ and real, then its eigenvalues and eigenvectors occur in complex conjugate pairs.

- We know eigenvalues must come in complex conjugate pairs because they are the roots of the characteristic polynomial with **real** coefficients

If A is $n \times n$ and real, then its eigenvalues and eigenvectors occur in complex conjugate pairs.

- We know eigenvalues must come in complex conjugate pairs because they are the roots of the characteristic polynomial with **real** coefficients

$$\overline{z_1 \cdot z_2} = \overline{z}_1 \cdot \overline{z}_2$$

- $\overline{A \cdot v} = \overline{A} \cdot \overline{v} = A \cdot \overline{v}$ (because A is real)
- $\overline{A \cdot v} = \overline{\lambda \cdot v} = \overline{\lambda} \cdot \overline{v}$ (because $Av = \lambda v$)

Eigenvalues of a real symmetric $n \times n$ matrix A are real

Eigenvalues of a real symmetric $n \times n$ matrix A are real

- **Definition:** A matrix A is **symmetric** if $A^T = A$

Eigenvalues of a real symmetric $n \times n$ matrix A are real

- **Definition:** A matrix A is **symmetric** if $A^T = A$
- **Proposition:** If A is real and symmetric, then $\forall\, x, y \in \mathbb{C}^n,\ \langle Ax, y\rangle = \langle x, Ay\rangle$

Eigenvalues of a real symmetric $n \times n$ matrix A are real

- **Definition:** A matrix A is **symmetric** if $A^T = A$
- **Proposition:** If A is real and symmetric, then $\forall\, x, y \in \mathbb{C}^n,\ \langle Ax, y \rangle = \langle x, Ay \rangle$
  - Choose this inner product: $< x, y >= x^T \overline{y}$

Eigenvalues of a real symmetric $n \times n$ matrix A are real

- **Definition:** A matrix A is **symmetric** if $A^T = A$
- **Proposition:** If A is real and symmetric, then $\forall\, x, y \in \mathbb{C}^n,\ \langle Ax, y\rangle = \langle x, Ay\rangle$
  - Choose this inner product: $< x, y >= x^T \overline{y}$

$$\langle Ax, y\rangle = x^\top A^\top \overline{y} = x^\top A \overline{y} \qquad \langle x, Ay\rangle = x^\top \overline{Ay} = x^\top \overline{A}\overline{y} = x^\top A \overline{y}$$

Eigenvalues of a real symmetric $n \times n$ matrix A are real

- **Definition:** A matrix A is **symmetric** if $A^T = A$
- **Proposition:** If A is real and symmetric, then $\forall\, x, y \in \mathbb{C}^n,\ \langle Ax, y\rangle = \langle x, Ay\rangle$
  - Choose this inner product: $< x, y >= x^T \overline{y}$

$$\langle Ax, y\rangle = x^\top A^\top \overline{y} = x^\top A \overline{y} \qquad \langle x, Ay\rangle = x^\top \overline{Ay} = x^\top \overline{A}\overline{y} = x^\top A \overline{y}$$

- **Proof:**

$$\langle Av, v\rangle = \langle v, Av\rangle$$
$$\Updownarrow$$
$$\langle \lambda v, v\rangle = \langle v, \lambda v\rangle$$
$$\Updownarrow$$
$$\lambda\langle v, v\rangle = \overline{\lambda}\langle v, v\rangle$$

Eigenvalues of a real symmetric $n \times n$ matrix A are real

- **Definition:** A matrix A is **symmetric** if $A^T = A$
- **Proposition:** If A is real and symmetric, then $\forall\, x, y \in \mathbb{C}^n,\ \langle Ax, y\rangle = \langle x, Ay\rangle$
  - Choose this inner product: $< x, y >= x^T \overline{y}$

$$\langle Ax, y\rangle = x^\top A^\top \overline{y} = x^\top A \overline{y} \qquad\qquad \langle x, Ay\rangle = x^\top \overline{Ay} = x^\top \overline{A}\overline{y} = x^\top A \overline{y}$$

- **Proof:**

$$\langle Av, v\rangle = \langle v, Av\rangle$$
$$\Updownarrow$$
$$\langle \lambda v, v\rangle = \langle v, \lambda v\rangle$$
$$\Updownarrow$$
$$\lambda\langle v, v\rangle = \overline{\lambda}\langle v, v\rangle$$

Note $<v, v> \neq 0$

Eigenvalues of a real symmetric $n \times n$ matrix A are real

- **Definition:** A matrix A is **symmetric** if $A^T = A$
- **Proposition:** If A is real and symmetric, then $\forall\, x, y \in \mathbb{C}^n,\ \langle Ax, y\rangle = \langle x, Ay\rangle$
  - Choose this inner product: $< x, y >= x^T \overline{y}$

$$\langle Ax, y\rangle = x^\top A^\top \overline{y} = x^\top A \overline{y} \qquad \langle x, Ay\rangle = x^\top \overline{Ay} = x^\top \overline{A}\overline{y} = x^\top A \overline{y}$$

- **Proof:**

$$\langle Av, v\rangle = \langle v, Av\rangle$$
$$\Updownarrow$$
$$\langle \lambda v, v\rangle = \langle v, \lambda v\rangle$$
$$\Updownarrow$$
$$\lambda \langle v, v\rangle = \overline{\lambda} \langle v, v\rangle$$

Then the corresponding eigenvector must be also real!

Note $<v, v> \neq 0$

Eigenvectors of a real symmetric $n \times n$ matrix A are orthogonal

Eigenvectors of a real symmetric $n \times n$ matrix A are orthogonal

- …is what we want to say. But that is not right because eigenvectors of repeated eigenvalues can be chosen to be non-orthogonal

~~Eigenvectors of a real symmetric~~ $\cancel{n \times n}$ ~~matrix~~ $\cancel{A}$ ~~are orthogonal~~

- …is what we want to say. But that is not right because eigenvectors of repeated eigenvalues can be chosen to be non-orthogonal
- Let A be an $n \times n$ real symmetric matrix and let λ_1 and λ_2 be distinct (real) eigenvalues. Then the corresponding (real) eigenvectors are orthogonal

~~Eigenvectors of a real symmetric $n \times n$ matrix A are orthogonal~~

- …is what we want to say. But that is not right because eigenvectors of repeated eigenvalues can be chosen to be non-orthogonal
- Let A be an $n \times n$ real symmetric matrix and let λ_1 and λ_2 be distinct (real) eigenvalues. Then the corresponding (real) eigenvectors are orthogonal
- **Proof:** Show that for the corresponding eigenvectors $\langle v^1, v^2 \rangle = 0$

$$\langle Av^1, v^2 \rangle = \langle v^1, Av^2 \rangle$$
$$\Updownarrow$$
$$\langle \lambda_1 v^1, v^2 \rangle = \langle v^1, \lambda_2 v^2 \rangle$$
$$\Updownarrow$$
$$\lambda_1 \langle v^1, v^2 \rangle = \lambda_2 \langle v^1, v^2 \rangle$$
$$\Updownarrow$$
$$0 = (\lambda_1 - \lambda_2) \langle v^1, v^2 \rangle$$

~~Eigenvectors of a real symmetric $n \times n$ matrix A are orthogonal~~

- …is what we want to say. But that is not right because eigenvectors of repeated eigenvalues can be chosen to be non-orthogonal
- Let A be an $n \times n$ real symmetric matrix and let λ_1 and λ_2 be distinct (real) eigenvalues. Then the corresponding (real) eigenvectors are orthogonal
- **Proof:** Show that for the corresponding eigenvectors $<v^1, v^2> = 0$

$$\langle Av^1, v^2 \rangle = \langle v^1, Av^2 \rangle$$
$$\Updownarrow$$
$$\langle \lambda_1 v^1, v^2 \rangle = \langle v^1, \lambda_2 v^2 \rangle$$
$$\Updownarrow$$
$$\lambda_1 \langle v^1, v^2 \rangle = \lambda_2 \langle v^1, v^2 \rangle$$
$$\Updownarrow$$
$$0 = (\lambda_1 - \lambda_2) \langle v^1, v^2 \rangle$$

Note $\lambda_1 - \lambda_2 \neq 0$

~~Eigenvectors of a real symmetric $n \times n$ matrix A are orthogonal~~

- The eigenvectors of an $n \times n$ real symmetric matrix can always be chosen to form an orthonormal basis of $\mathbb{R}^n$

~~Eigenvectors of a real symmetric $n \times n$ matrix A are orthogonal~~

- The eigenvectors of an $n \times n$ real symmetric matrix can always be chosen to form an orthonormal basis of $\mathbb{R}^n$
- Repeated eigenvalues require special consideration (deflected to homework exercise)

~~Eigenvectors of a real symmetric $n \times n$ matrix A are orthogonal~~

- The eigenvectors of an $n \times n$ real symmetric matrix can always be chosen to form an orthonormal basis of $\mathbb{R}^n$
- Repeated eigenvalues require special consideration (deflected to homework exercise)
- Consider a real but not symmetric matrix

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

Definition: Orthogonal Matrix

- An $n \times n$ real matrix is **orthogonal** if $Q^T Q = I$

Definition: Orthogonal Matrix

- An $n \times n$ real matrix is **orthogonal** if $Q^T Q = I$

- Remark: $Q^T = Q^{-1}$

Definition: Orthogonal Matrix

- An $n \times n$ real matrix is **orthogonal** if $Q^T Q = I$

- Remark: $Q^T = Q^{-1}$
- Orthogonal because

$$\langle Q_i, Q_j \rangle = Q_i^\top Q_j =: [Q^\top Q]_{ij} = [I]_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Definition: Orthogonal Matrix

- An $n \times n$ real matrix is **orthogonal** if $Q^T Q = I$

- Remark: $Q^T = Q^{-1}$
- Orthogonal because

$$\langle Q_i, Q_j \rangle = Q_i^\top Q_j =: [Q^\top Q]_{ij} = [I]_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

- Column vectors of an orthogonal matrix form an orthonormal basis of $\mathbb{R}^n$

Diagonalization with Orthogonal Matrix

- Suppose A is an $n \times n$ real symmetric matrix. Then there exists an orthogonal matrix Q such that

$$Q^T AQ = \Lambda = \mathrm{diag}\{\lambda_1, \ldots, \lambda_n\}$$

Diagonalization with Orthogonal Matrix

- Suppose A is an $n \times n$ real symmetric matrix. Then there exists an orthogonal matrix Q such that

$$Q^T A Q = \Lambda = \mathrm{diag}\{\lambda_1, \ldots, \lambda_n\}$$

- **Proof Idea:**

$$Av = \lambda v$$

Diagonalization with Orthogonal Matrix

- Suppose A is an $n \times n$ real symmetric matrix. Then there exists an orthogonal matrix Q such that

$$Q^T AQ = \Lambda = \mathrm{diag}\{\lambda_1, ..., \lambda_n\}$$

- **Proof Idea:**

$$Av = \lambda v$$

- Very useful property for later (e.g., **singular value decomposition**)

Orthogonal Matrices Preserve Norms

- Norm-preserving

$$\|Qx\| = \|x\|$$

Orthogonal Matrices Preserve Norms

- Norm-preserving

$$\|Qx\| = \|x\|$$

$$\|Qx\|^2 = (Qx)^T Qx = x^T Q^T Qx = x^T (I) x = x^T x = \|x\|^2$$

- Rotations and reflections preserve length

Eigenvalues of A^TA and AA^T are non-negative (real numbers)

Eigenvalues of A^TA and AA^T are non-negative (real numbers)

- **Proof for** A^TA:

Eigenvalues of A^TA and AA^T are non-negative (real numbers)

- **Proof for A^TA:**
  - Let $A^TAv = \lambda v$ where $v \in \mathbb{R}^n$, $v \neq 0$, $\lambda \in \mathbb{R}$.

Eigenvalues of A^TA and AA^T are non-negative (real numbers)

- **Proof for A^TA:**
  - Let $A^TAv = \lambda v$ where $v \in \mathbb{R}^n$, $v \neq 0$, $\lambda \in \mathbb{R}$.
  - Show: $\lambda \geq 0$

Eigenvalues of $A^T A$ and AA^T are non-negative (real numbers)

- **Proof for $A^T A$:**
  - Let $A^T A v = \lambda v$ where $v \in \mathbb{R}^n$, $v \neq 0$, $\lambda \in \mathbb{R}$.

$$v^\top A^\top A v = v^\top \lambda v$$
$$\Updownarrow$$
$$\langle Av, Av \rangle = \lambda \langle v, v \rangle$$
$$\Updownarrow$$
$$\|Av\|^2 = \lambda \|v\|^2.$$

  - Show: $\lambda \geq 0$

Quadratic Forms and Positive (Semi)Definiteness

Definition: Quadratic Forms

- Let M be an $n \times n$ real matrix and $x \in \mathbb{R}^n$. Then $x^T M x$ is called a **quadratic form**

Definition: Quadratic Forms

- Let M be an $n \times n$ real matrix and $x \in \mathbb{R}^n$. Then $x^T M x$ is called a **quadratic form**
  - Least squares $<x, x> = x^T x$
    - Quadratic form where $M = I$

Definition: Quadratic Forms

- Let M be an $n \times n$ real matrix and $x \in \mathbb{R}^n$. Then x^TMx is called a **quadratic form**
    - Least squares $< x, x > = x^T x$
        - Quadratic form where $M = I$
    - Weighted least squares $< x, x > = x^T M x$
        - In an estimation problem, not all components of error should be equally weighted! Depends on the certainty of the measurements. Adjusting M allows for this flexibility

Definition: Quadratic Forms

- Let M be an $n \times n$ real matrix and $x \in \mathbb{R}^n$. Then $x^T M x$ is called a **quadratic form**
  - Least squares $< x, x > = x^T x$
    - Quadratic form where $M = I$
  - Weighted least squares $< x, x > = x^T M x$
    - In an estimation problem, not all components of error should be equally weighted! Depends on the certainty of the measurements. Adjusting M allows for this flexibility
    - M cannot be chosen randomly! Otherwise, no longer inner product/norm

Definition: Quadratic Forms

- Let M be an $n \times n$ real matrix and $x \in \mathbb{R}^n$. Then x^TMx is called a **quadratic form**
  - Least squares $< x, x > = x^T x$
    - Quadratic form where $M = I$
  - Weighted least squares $< x, x > = x^T M x$
    - In an estimation problem, not all components of error should be equally weighted! Depends on the certainty of the measurements. Adjusting M allows for this flexibility
    - M cannot be chosen randomly! Otherwise, no longer inner product/norm

Symmetric Part of Quadratic Form

- Any square matrix M has a symmetric and skew-symmetric part

$$M = \underbrace{\frac{M + M^\top}{2}}_{symmetric} + \underbrace{\frac{M - M^\top}{2}}_{skew\ symmetric}$$

Symmetric Part of Quadratic Form

- Any square matrix M has a symmetric and skew-symmetric part

$$M = \underbrace{\frac{M + M^\top}{2}}_{symmetric} + \underbrace{\frac{M - M^\top}{2}}_{skew\ symmetric}$$

- For the quadratic form $x^T M x$ only the symmetric part matters. Thus, all theorems about quadratic forms consider only symmetric M.

$$x^\top M x = x^\top \left(\frac{M+M^\top}{2}\right) x$$

Connections with Eigenvalues and Quadratic Forms

- For a real symmetric matrix form M. Then $\forall\, x \in \mathbb{R}^n$

$$\lambda_{\min}\, x^\top x \leq x^\top M x \leq \lambda_{\max}\, x^\top x$$

- Bounds are “tight”

Theorem: Positive Definite

- **Definition:** A real symmetric matrix form P is **positive definite** if

$$\forall x \in \mathbb{R}^n, x \neq 0 \quad \Rightarrow \quad x^T P x > 0$$

- **Theorem:** A symmetric real matrix P is **positive definite** if and only if all of its eigenvalues are **strictly greater than** zero.

- **Notation:** $P > 0$

Theorem: Positive Semidefinite

- **Definition:** A real symmetric matrix form P is **positive semidefinite** if

$$\forall x \in \mathbb{R}^n, x \neq 0 \quad \Rightarrow \quad x^T P x \geq 0$$

- **Theorem:** A symmetric real matrix P is **positive semidefinite** if and only if all of its eigenvalues are **greater than or equal** to zero.

- **Notation:** $P \geq 0$

Fun fact: Square Root of a Matrix

- **Definition:** N is a square root of a real symmetric matrix P if $N^T N = P$.

Fun fact: Square Root of a Matrix

- **Definition:** N is a square root of a real symmetric matrix P if $N^T N = P$.
  - We note that $N^T N = (N^T N)^T$ (is always symmetric)
  - There are several notions of what "square root" means for a matrix. This is the simplest one

Fun fact: Square Root of a Matrix

- **Definition:** N is a square root of a real symmetric matrix P if $N^T N = P$.
  - We note that $N^T N = (N^T N)^T$ (is always symmetric)
  - There are several notions of what "square root" means for a matrix. This is the simplest one
- **Theorem:** $P \geq 0$ if, and only if, $\exists N$ such that $N^T N = P$.

Fun fact: Square Root of a Matrix

- **Definition:** N is a square root of a real symmetric matrix P if $N^T N = P$.
  - We note that $N^T N = (N^T N)^T$ (is always symmetric)
  - There are several notions of what "square root" means for a matrix. This is the simplest one
- **Theorem:** $P \geq 0$ if, and only if, $\exists N$ such that $N^T N = P$.
- **Proof Outline:**

$$P = O^T \Lambda O$$

Square root $N = \Lambda^{1/2} O$

Fun fact: Square Root of a Matrix

- **Definition:** N is a square root of a real symmetric matrix P if $N^TN = P$.
  - We note that $N^TN = (N^TN)^T$ (is always symmetric)
  - There are several notions of what "square root" means for a matrix. This is the simplest one
- **Theorem:** $P \geq 0$ if, and only if, $\exists N$ such that $N^TN = P$.
- **Proof Outline:**

$$P = O^T \Lambda O$$

$$\Lambda^{1/2} := \mathrm{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n}),$$

Square root $N = \Lambda^{1/2} O$

$$(\Lambda^{1/2})^\top \Lambda^{1/2} = \Lambda^{1/2}\Lambda^{1/2} = \Lambda.$$

$$N^\top N = O^\top \left(\Lambda^{1/2}\right)^\top \Lambda^{1/2} O = O^\top \Lambda O = P.$$

Positive-Definiteness Checks

- Positive definite matrices have many interesting and useful properties
  - Deep significance in relation to quadratic forms

Positive-Definiteness Checks

- Positive definite matrices have many interesting and useful properties
  - Deep significance in relation to quadratic forms
- Check whether a certain matrix is positive definite:

Positive-Definiteness Checks

- Positive definite matrices have many interesting and useful properties
  - Deep significance in relation to quadratic forms
- Check whether a certain matrix is positive definite:
  - Find eigenvalues. **HARD**
  - **You should only look for things that you have to**
  - **For example, you don’t have to invert a matrix to check if the column vectors are linearly independent**

Theorem: Schur's Complements

- Suppose that A, B and C are real matrices, $A = n \times n$ is symmetric, $B = n \times m$, and $C = m \times m$ is symmetric.

Then for the symmetric matrix $$M = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$$

The following are equivalent

(a) $M > 0.$

(b) $A > 0,$ *and* $C - B^\top A^{-1} B > 0.$

(c) $C > 0,$ *and* $A - BC^{-1} B^\top > 0.$

Theorem: Schur's Complements

- Suppose that A, B and C are real matrices, $A = n \times n$ is symmetric, $B = n \times m$, and $C = m \times m$ is symmetric.

Then for the symmetric matrix $$M = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$$

The following are equivalent

$(a)\ \ M > 0.$

Schur's Complement of A in M

$(b)\ \ A > 0,\ and\ \boxed{C - B^\top A^{-1} B} > 0.$

$(c)\ \ C > 0,\ and\ \boxed{A - BC^{-1} B^\top} > 0.$

Schur's Complement of C in M

Theorem: Schur's Complements

- Suppose that A, B and C are real matrices, $A = n \times n$ is symmetric, $B = n \times m$, and $C = m \times m$ is symmetric.

Then for the symmetric matrix $$M = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \quad M_1 = \begin{bmatrix} 3 & -2 \\ -2 & 3 \end{bmatrix}$$

The following are equivalent

(a) $M > 0.$

Schur's Complement of A in M

(b) $A > 0,$ *and* $\boxed{C - B^\top A^{-1} B} > 0.$

(c) $C > 0,$ *and* $\boxed{A - BC^{-1} B^\top} > 0.$

Schur's Complement of C in M

Theorem: Schur's Complements

- Suppose that A, B and C are real matrices, $A = n \times n$ is symmetric, $B = n \times m$, and $C = m \times m$ is symmetric.

Then for the symmetric matrix $$M = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \quad M_1 = \left[\begin{array}{c|c} 3 & -2 \\ \hline -2 & 3 \end{array}\right]$$

The following are equivalent

$(a)\ \ M > 0.$

Schur's Complement of A in M

$(b)\ \ A > 0, \textit{ and } \boxed{C - B^\top A^{-1} B} > 0.$

$(c)\ \ C > 0, \textit{ and } \boxed{A - BC^{-1} B^\top} > 0.$

Schur's Complement of C in M

Theorem: Schur's Complements

- Suppose that A, B and C are real matrices, $A = n \times n$ is symmetric, $B = n \times m$, and $C = m \times m$ is symmetric.

Then for the symmetric matrix $$M = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \quad M_3 = \begin{bmatrix} \alpha & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 3 \end{bmatrix}$$

The following are equivalent

$(a)\ M > 0.$

Schur's Complement of A in M

$(b)\ A > 0, \textit{and } \boxed{C - B^\top A^{-1} B} > 0.$

$(c)\ C > 0, \textit{and } \boxed{A - BC^{-1} B^\top} > 0.$

Schur's Complement of C in M

Theorem: Schur's Complements

- Suppose that A, B and C are real matrices, $A = n \times n$ is symmetric, $B = n \times m$, and $C = m \times m$ is symmetric.

Then for the symmetric matrix $$M = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \quad M_3 = \left[\begin{array}{c|cc} \alpha & 1 & 1 \\ \hline 1 & 2 & 1 \\ 1 & 1 & 3 \end{array}\right]$$

The following are equivalent

$(a)\ \ M > 0.$

Schur's Complement of A in M

$(b)\ \ A > 0, \textit{ and } \boxed{C - B^\top A^{-1} B} > 0.$

$(c)\ \ C > 0, \textit{ and } \boxed{A - BC^{-1} B^\top} > 0.$

Schur's Complement of C in M