6.4 Minimum principles. We exploit simple calculus:

The function $P(x) = \frac{1}{2}ax^2 - bx$ is stationary when $P'(x)=0$: $ax=b$

Here we take a matrix formulation, and rephrase it so that instead of stationary we want a minimum — so that now we need $a>0$. This explains our formulation here:

Principle. Let A be symmetric positive definite. Then

$$P(x) = \frac{1}{2}x^TAx - x^Tb$$

has a minimum where $Ax=b$. So at that point,

$$P = -\frac{1}{2}b^TA^{-1}b.$$

Proof. Suppose $Ax=b$, and let y be any vector. Then

$$P(y) - P(x) = \frac{1}{2}y^TAy - y^Tb - \frac{1}{2}x^TAx + x^Tb$$

But $Ax=b$, so this becomes

$$= \frac{1}{2}y^TAy - y^TAx + \frac{1}{2}x^TAx$$

$$= \frac{1}{2}(y-x)^TA(y-x)$$ Now use that A is p.d.!

Back to least squares: $\|Ax-b\|$ minimized at $x=\hat{x}$.

But

$$\|Ax-b\|^2 = (Ax-b)^T(Ax-b) = (x^TA^T-b^T)(Ax-b)$$

$$= \underbrace{x^TA^TAx}_{\text{scalar}} - 2\underbrace{x^TA^Tb}_{\text{scalar}} + b^Tb.$$

This has form $2\left[\frac{1}{2}x^T\boxed{A^TA}x - x^T\boxed{A^Tb}\right] + $ const

(Remember, p 162)

$$\hat{x}: A^TA\hat{x} = A^Tb \text{ or}$$

$$\boxed{\hat{x} = (A^TA)^{-1}A^Tb}$$

Rayleigh quotient Eigenvalues connected to minimization — but be aware whether A is to be P.D or not, or symmetric.

Rayleigh quotient: Smallest eigenvalue λ_1 of A is

$$R(x_1) = \min_{|x|=1} \frac{x^TAx}{x^Tx}$$ (book sh say A is symmetric)

Maximum is $\lambda_n = \max_{|x|=1} \frac{x^TAx}{x^Tx}$

(By choice of basis, $x^TAx = \lambda_1 x_1 x_1^T$ ... $\lambda_n x^T x_n$.

P 345 #6 $A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$

So $\det A - \lambda I = \lambda^2 - 4\lambda + 3 = (\lambda - 3)(\lambda - 1)$, Evectors:

$\lambda = 3$: $\frac{1}{\sqrt{2}}(1, -1)$; $\lambda = 1$: $\frac{1}{\sqrt{2}}(1, 1)$ so we have

$2x^2 - 2xy + 2y^2 = 3\left(\frac{x}{\sqrt{2}} - \frac{y}{\sqrt{2}}\right) + 1\left(\frac{x}{\sqrt{2}} + \frac{y}{\sqrt{2}}\right)^2$.

So minimum value with $|x| = 1$ comes at $(1/\sqrt{2}, 1/\sqrt{2})$.

Note #11 $R(x) = \dfrac{\lambda_1 y_1^2 + \cdots + \lambda_n y_n^2}{y_1^2 + y_2^2 + \cdots + y_n^2}$ with $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$,

<u>Chapter 7</u> Matrix norm, stability <u>Norm of a matrix</u>.

<u>Def</u> $\|A\| = \max\limits_{\|x\|=1} \dfrac{\|Ax\|}{\|x\|} = \max\limits_{\|x\|=1} \|Ax\|$.

($\|x\| = 1$ is simple normalization). Note: $\|Ax\| \le \|A\| \, \|x\|$

For p.d. This is easy; it is the <u>largest eigenvalue</u>.

But that's not so clear in general. It is related to a question that is of obvious practical value;

Let $Ax = b$,

so that $\|x\| \ge \|b\|, \lambda_{max}$

but replace b by $b' = b + \delta b$ with δb small, We'd like to say we don't change x by much, but that just isn't true

<u>If</u> we define δx so that

$$A(x + \delta x) = b + \delta b, \quad \text{or } A(\delta x) = \delta b,$$

Then

(*) $\delta x = A^{-1}(\delta b)$

So if A is symmetric p.d. with eigenvalues $0 < \lambda_1 \le \cdots \le \lambda_n$ then

$$|\delta x| \le \frac{\delta b}{\lambda_1} \leftarrow \text{smallest} \; ; \quad \frac{|\delta x|}{|x|} \le \frac{\lambda_{max}}{\lambda_{min}} \frac{|\delta b|}{|b|}$$

We want these in scale-invariant way — so ~~if we~~ multiply all entries by a constant we don't change anything

Thus we define the <u>condition #</u> of A (p.d.) as

$\dfrac{\lambda_{max}}{\lambda_{min}}$

Note: $\begin{pmatrix} \varepsilon & 0 \\ 0 & \varepsilon \end{pmatrix}$ has small determinant, optimal condition #

6.4 Minimum principles. We exploit simple calculus:

The function $P(x) = \frac{1}{2}ax^2 - bx$ is stationary when $P'(x) = 0$: $ax = b$

Here we take a matrix formulation, and rephrase it so that instead of stationary we want a minimum — so that now we need $a > 0$. This explains our formulation here:

Principle. Let A be symmetric positive definite. Then

$$P(x) = \frac{1}{2}x^T A x - x^T b$$

has a minimum where $Ax = b$. So at that point,

$$P = -\frac{1}{2}b^T A^{-1} b.$$

Proof. Suppose $Ax = b$, and let y be any vector. Then

$$P(y) - P(x) = \frac{1}{2}y^T A y - y^T b - \frac{1}{2}x^T A x + x^T b$$

But $Ax = b$, so this becomes

$$= \frac{1}{2}y^T A y - y^T A x + \frac{1}{2}x^T A x$$

$$= \frac{1}{2}(y - x)^T A (y - x) \quad \text{Now use that } A \text{ is p.d.!}$$

Back to least squares: $\|Ax - b\|$ minimized at $x = \hat{x}$.

But

$$\|Ax - b\|^2 = (Ax - b)^T(Ax - b) = (x^T A^T - b^T)(Ax - b)$$

$$= \underbrace{x^T A^T A x}_{\text{scalar}} - 2\underbrace{x^T A^T b}_{\text{scalar}} + b^T b.$$

This has form

$$2\left[\frac{1}{2}x^T \boxed{A^T A} x - x^T \boxed{A^T b}\right] + \text{const}$$

(Remember, p 162)

$$\hat{x}: A^T A \hat{x} = A^T b \text{ or}$$

$$\boxed{\hat{x} = (A^T A)^{-1} A^T b}$$

Rayleigh quotient. Eigenvalues connected to minimization — but be aware whether A is to be P.D or not, or symmetric.

Rayleigh quotient: Smallest eigenvalue λ_1 of A is

$$R(x_1) = \min_{|x|=1} \frac{x^T A x}{x^T x} \quad \text{(book sh. say } A \text{ is symmetric)}$$

Maximum is $\lambda_n = \max_{|x|=1} \frac{x^T A x}{x^T x}$.

(By choice of basis, $x^T A x = \lambda_1 x_1 x_1^T \cdots \lambda_n x_n^T x_n$.