Table 8.2. Pressure (in mm of mercury) and temperature (°C) of a gas at constant volume.

Trial number i	"x" Pressure P_i	"y" Temperature T_i	$A + BP_i$
1	65	−20	−22.2
2	75	17	14.9
3	85	42	52.0
4	95	94	89.1
5	105	127	126.2

Linear Least Squares

Can also talk about non-square systems where
A is $m\times n$, b is $m\times 1$, and x is $n\times 1$

- *Overdetermined* if $m>n$:
  "more equations than unknowns"
- *Underdetermined* if $n>m$:
  "more unknowns than equations"
  Can look for best solution using least squares

Better written $\boldsymbol{Ax} \cong \boldsymbol{b}$, since equality usually
not exactly satisfiable when $m > n$

Linear Least Squares

Can also talk about non-square systems where
A is $m \times n$, b is $m \times 1$, and x is $n \times 1$

- *Overdetermined* if $m>n$:
  "more equations than unknowns"
- *Underdetermined* if $n>m$:
  "more unknowns than equations"
  Can look for best solution using least squares

Better written $\boldsymbol{Ax} \cong \boldsymbol{b}$, since equality usually
not exactly satisfiable when $m > n$

Least squares solution $\boldsymbol{x}$ minimizes squared Euclidean norm of residual vector $\boldsymbol{r} = \boldsymbol{b} - \boldsymbol{Ax}$,

$$\min_{\boldsymbol{x}} \|\boldsymbol{r}\|_2^2 = \min_{\boldsymbol{x}} \|\boldsymbol{b} - \boldsymbol{Ax}\|_2^2$$

Normal Equations

Least squares minimizes squared Euclidean norm

$$\|\boldsymbol{r}\|_2^2 = \boldsymbol{r}^T\boldsymbol{r}$$

of residual vector

$$\boldsymbol{r} = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}$$

Minimization Technique?

Normal Equations

Least squares minimizes squared Euclidean norm

$$\|r\|_2^2 = r^T r$$

of residual vector

$$r = b - Ax$$

Minimization Technique

Step 1

$$\begin{aligned}\|r\|_2^2 &= r^T r = (b - Ax)^T (b - Ax) \\ &= b^T b - 2x^T A^T b + x^T A^T A x,\end{aligned}$$

Normal Equations

Least squares minimizes squared Euclidean norm

$$\|\boldsymbol{r}\|_2^2 = \boldsymbol{r}^T\boldsymbol{r}$$

of residual vector

$$\boldsymbol{r} = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}$$

Minimization Technique

Step 1

$$\phi = \| r\|_2 = r^T r = (\boldsymbol{b} - \boldsymbol{Ax})^T(\boldsymbol{b} - \boldsymbol{Ax})$$
$$= \boldsymbol{b}^T\boldsymbol{b} - 2\boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{b} + \boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{Ax}$$

Step 2

The necessary condition for minimum is to obtain a critical point of ϕ

i.e $\nabla\phi(x) = 0$ (i th component $\frac{\partial\phi(x)}{\partial x_i}$)

Normal Equations

Least squares minimizes squared Euclidean norm

$$\|r\|_2^2 = r^T r$$

of residual vector

$$r = b - Ax$$

Minimization Technique

Step 1

$$\phi = \| r\|_2 = r^T r = (\boldsymbol{b} - \boldsymbol{Ax})^{\boldsymbol{T}}(\boldsymbol{b} - \boldsymbol{Ax})$$
$$= \boldsymbol{b}^{\boldsymbol{T}}\boldsymbol{b} - \boldsymbol{2x}^{\boldsymbol{T}}\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{b} + \boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{Ax}$$

Step 2

The necessary condition for minimum is to obtain a critical point of ϕ

i.e $\nabla\phi(x) = 0$ (i th component $\frac{\partial\phi(x)}{\partial x_i}$)

$$\phi = \| r\|_2 = r^T r = (\boldsymbol{b} - \boldsymbol{Ax})^{\boldsymbol{T}}(\boldsymbol{b} - \boldsymbol{Ax})$$
$$0 = -\boldsymbol{2A}^{\boldsymbol{T}}\boldsymbol{b} + \boldsymbol{2A}^{\boldsymbol{T}}\boldsymbol{Ax}$$
$$\therefore \boldsymbol{x} = \left(\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A}\right)^{-\boldsymbol{1}}\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{b}$$

Normal Equations

Least squares minimizes squared Euclidean norm

$$\|r\|_2^2 = r^T r$$

of residual vector

$$r = b - Ax$$

Minimization Technique

Step 1

$$\phi = \| r\|_2 = r^T r = (\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x})^T(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x})$$
$$= \boldsymbol{b}^T\boldsymbol{b} - \boldsymbol{2}\boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{b} + \boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x}$$

Step 2

The necessary condition for minimum is to obtain a critical point of ϕ

i.e $\nabla\phi(x) = 0$ (i th component $\frac{\partial\phi(x)}{\partial x_i}$)

$$\phi = \| r\|_2 = r^T r = (\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x})^T(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x})$$
$$\nabla\phi(x) = 0 \quad = -\boldsymbol{2}\boldsymbol{A}^T\boldsymbol{b} + \boldsymbol{2}\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x}$$
$$\therefore \boldsymbol{x} = \left(\boldsymbol{A}^T\boldsymbol{A}\right)^{-1}\boldsymbol{A}^T\boldsymbol{b}$$

Sufficient condition: Hessian matrix of second partial derivative which is $2A^T A$ is positive definite.

Normal Equations

Least squares minimizes squared Euclidean norm

$$\|r\|_2^2 = r^T r$$

of residual vector

$$r = b - Ax$$

Minimization Technique

Step 1

$$\phi = \| r \|_2 = r^T r = (\boldsymbol{b} - \boldsymbol{Ax})^{\boldsymbol{T}} (\boldsymbol{b} - \boldsymbol{Ax})$$
$$= \boldsymbol{b^T b} - \boldsymbol{2x^T A^T b} + \boldsymbol{x^T A^T A x}$$

Step 2

The necessary condition for minimum is to obtain a critical point of ϕ

i.e $\nabla\phi(x) = 0$ (i th component $\frac{\partial\phi(x)}{\partial x_i}$)

$$\phi = \| r \|_2 = r^T r = (\boldsymbol{b} - \boldsymbol{Ax})^{\boldsymbol{T}} (\boldsymbol{b} - \boldsymbol{Ax})$$
$$\nabla\phi(x) = 0 = -\boldsymbol{2A^T b} + \boldsymbol{2A^T A x}$$
$$\therefore \boldsymbol{x} = \left(\boldsymbol{A^T A}\right)^{-1} \boldsymbol{A^T b}$$

Sufficient condition: Hessian matrix of second partial derivative which is $2A^T A$ is positive definite.

Possible if the columns of A are linearly dependent

$rank\ (A) = n$

Normal Equations

$2A^T b = 2A^T Ax$ Is commonly known as **normal equations**

Example: Pressure vs Temp

- Data is Overdetermined. Why?
- Five experiments, Two unknowns

Table 8.2. Pressure (in mm of mercury) and temperature (°C) of a gas at constant volume.

Trial number i	"x" Pressure P_i	"y" Temperature T_i	$A + BP_i$
1	65	−20	−22.2
2	75	17	14.9
3	85	42	52.0
4	95	94	89.1
5	105	127	126.2

Measurement of absolute zero with a constant-volume gas thermometer

Normal Equations

Example: Pressure vs Temp

Table 8.2. Pressure (in mm of mercury) and temperature (°C) of a gas at constant volume.

Trial number i	"x" Pressure P_i	"y" Temperature T_i	$A + BP_i$
1	65	−20	−22.2
2	75	17	14.9
3	85	42	52.0
4	95	94	89.1
5	105	127	126.2

Measurement of absolute zero with a constant-volume gas thermometer

- Assume Temp is linear function of Pressure : $T = A + BP$
- Q: What is **A?**

$$a + bP_1 \cong T_1$$

...

...

$$a + bP_5 \cong T_5$$

Normal Equations

Example: Pressure vs Temp

Table 8.2. Pressure (in mm of mercury) and temperature (°C) of a gas at constant volume.

Trial number i	"x" Pressure P_i	"y" Temperature T_i	$A + BP_i$
1	65	−20	−22.2
2	75	17	14.9
3	85	42	52.0
4	95	94	89.1
5	105	127	126.2

$$a + bP_1 \cong T_1$$

$$\dots$$

$$\dots$$

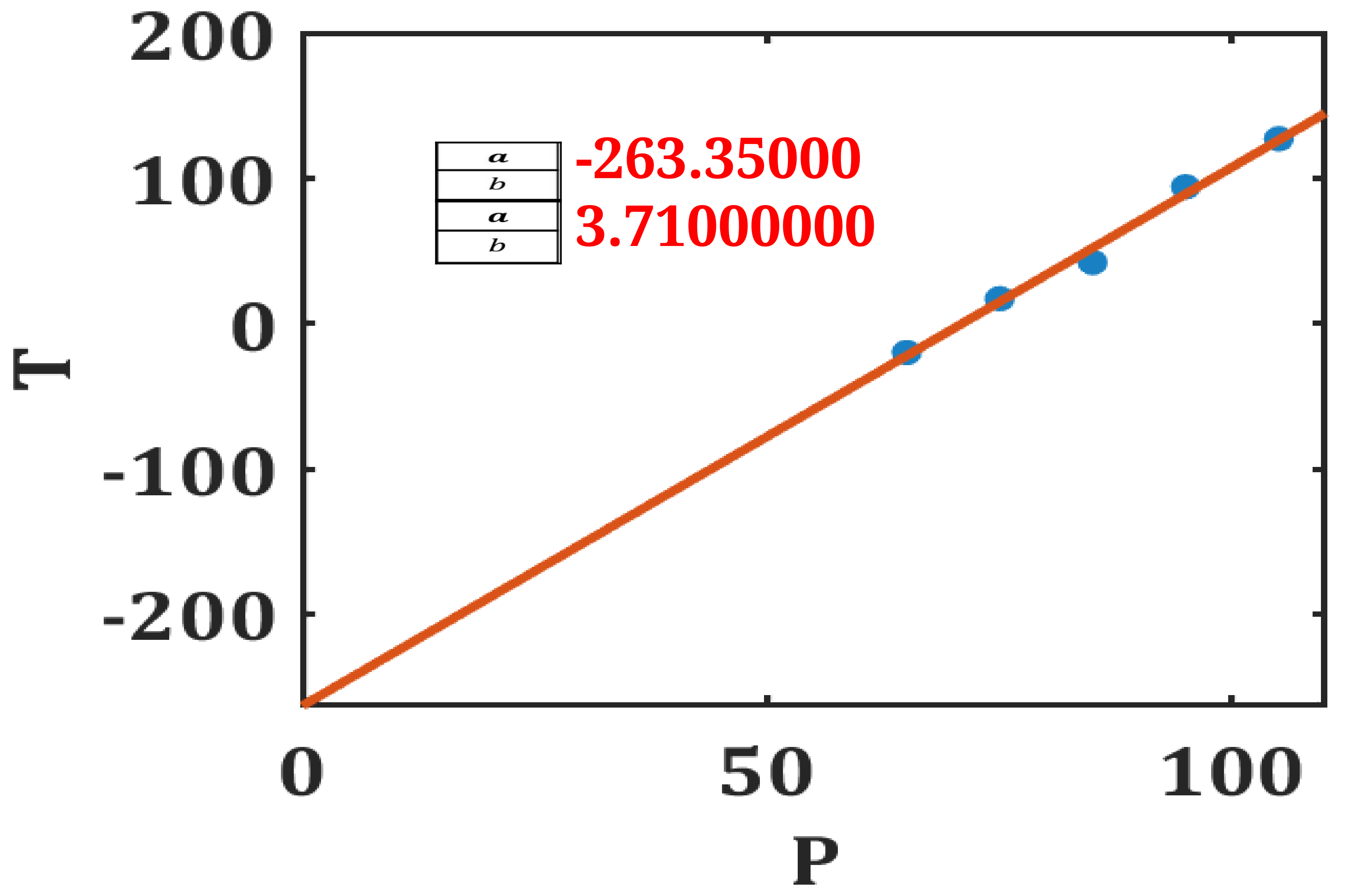
$$a + bP_5 \cong T_5$$

Normal Equations

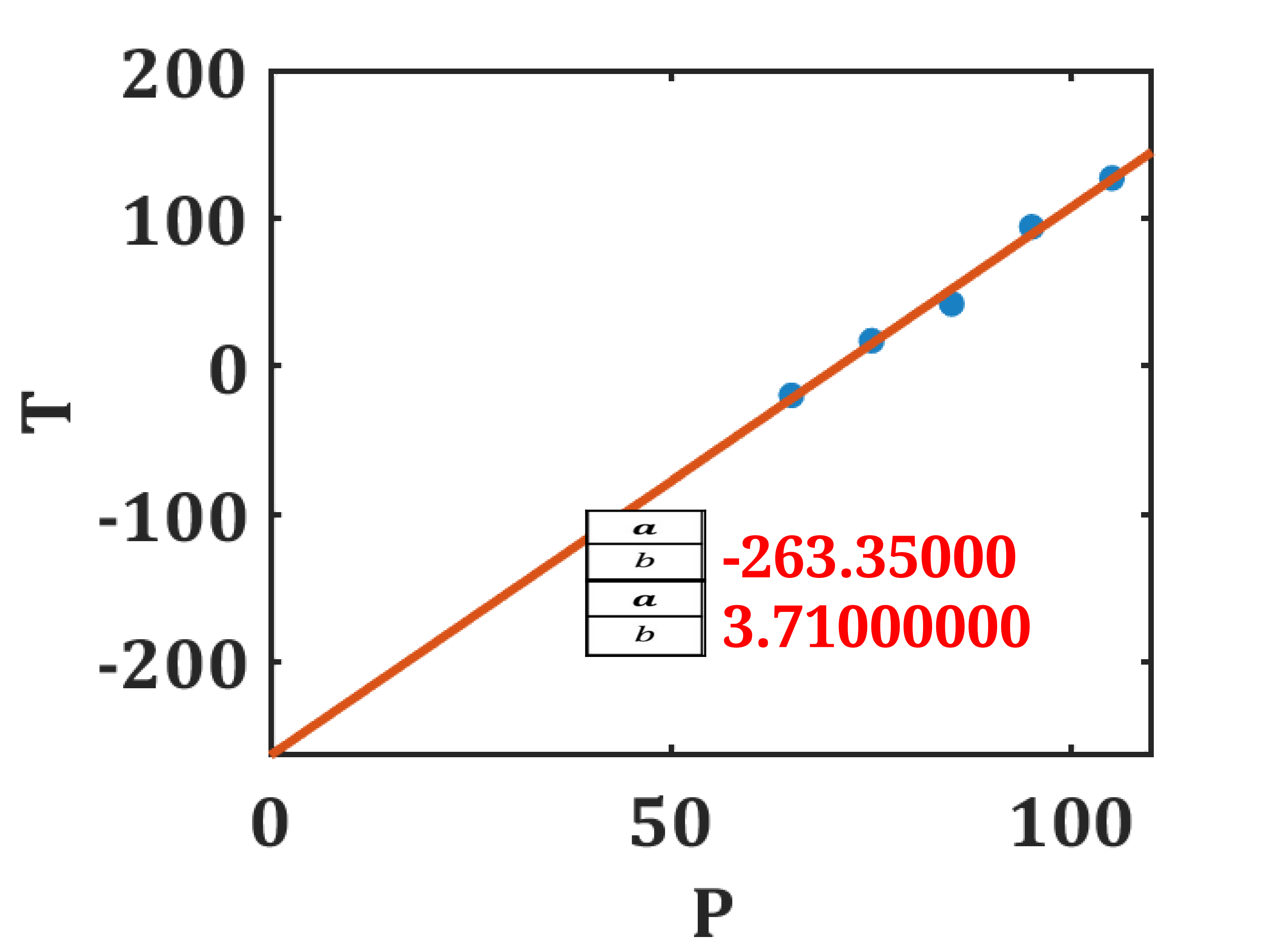
Example

Table 8.2.
(°C) of a g

Trial numb i
1
2
3
4
5



Normal Equations




```
clear all
%%% Temprature vs Pressure%%
%%% Given : T=a+bP
%%%% Find, a,b


A=[65 75 85 95 105];
b=[-20 17 42 94 127]';


%%% Ax=b %%%%
%XX=A;
XX=[ones(length(A),1) A'];
YY=b;

x=(inv(XX'*XX))*XX'*YY;


A1=0:10:111;
y=x(2,:)*A1 +x(1,:);
scatter(A,b,'filled','MarkerFaceAlpha',0.9, 'SizeData',75);
hold on
plot(A1,y,'linewidth',3 )
xlabel('P');
ylabel('T');
ylim=([-300 100]);
set(findall(gcf,'-property','FontSize'),'FontName','Cambria',...
'FontSize',24,'linewidth',2.0,'fontweight','b')
set(gcf,'InvertHardCopy','off','Color','white');
box on;
```

Normal Equations

$$2A^T b = 2A^T A x$$

Is commonly known as **normal equations**

Example: Polynomial fitting

- System can be treated as linear, Why?
- Overdetermined. Why?
- Five experiments, Three unknowns

- Matrix with columns (or rows) successive powers of independent variable called

$$x_1 + t_1 x_2 + t_1^2 x_3 \cong y_1$$
$$x_1 + t_2 x_2 + t_2^2 x_3 \cong y_2$$
$$\dots$$
$$\dots$$
$$\dots$$
$$x_1 + t_5 x_2 + t_5^2 x_3 \cong y_5$$

$$\boldsymbol{Ax} = \begin{bmatrix} 1 & t_1 & t_1^2 \\ 1 & t_2 & t_2^2 \\ 1 & t_3 & t_3^2 \\ 1 & t_4 & t_4^2 \\ 1 & t_5 & t_5^2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \cong \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} = \boldsymbol{b}$$

$$2A^T b = 2A^T Ax$$

Solving LLS problems

Example: Polynomial fitting

- System can be treated as linear, Why?
- Overdetermined. Why?
- Five experiments, Three unknowns

❑ Matrix with columns (or rows) successive powers of independent variable called Vandermonde matrix

$$\boldsymbol{Ax} = \begin{bmatrix} 1 & t_1 & t_1^2 \\ 1 & t_2 & t_2^2 \\ 1 & t_3 & t_3^2 \\ 1 & t_4 & t_4^2 \\ 1 & t_5 & t_5^2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \cong \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} = \boldsymbol{b}$$

t	-1.0	-0.5	0.0	0.5	1.0
y	1.0	0.5	0.0	0.5	2.0

$$\boldsymbol{Ax} = \begin{bmatrix} 1 & -1.0 & 1.0 \\ 1 & -0.5 & 0.25 \\ 1 & 0.0 & 0.0 \\ 1 & 0.5 & 0.25 \\ 1 & 1.0 & 1.0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \cong \begin{bmatrix} 1.0 \\ 0.5 \\ 0.0 \\ 0.5 \\ 2.0 \end{bmatrix} = \boldsymbol{b}$$

Solving LLS problems

$$2A^Tb = 2A^TAx$$

t	-1.0	-0.5	0.0	0.5	1.0
y	1.0	0.5	0.0	0.5	2.0

$$Ax = \begin{bmatrix} 1 & -1.0 & 1.0 \\ 1 & -0.5 & 0.25 \\ 1 & 0.0 & 0.0 \\ 1 & 0.5 & 0.25 \\ 1 & 1.0 & 1.0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \cong \begin{bmatrix} 1.0 \\ 0.5 \\ 0.0 \\ 0.5 \\ 2.0 \end{bmatrix} = b$$

LLS and Example

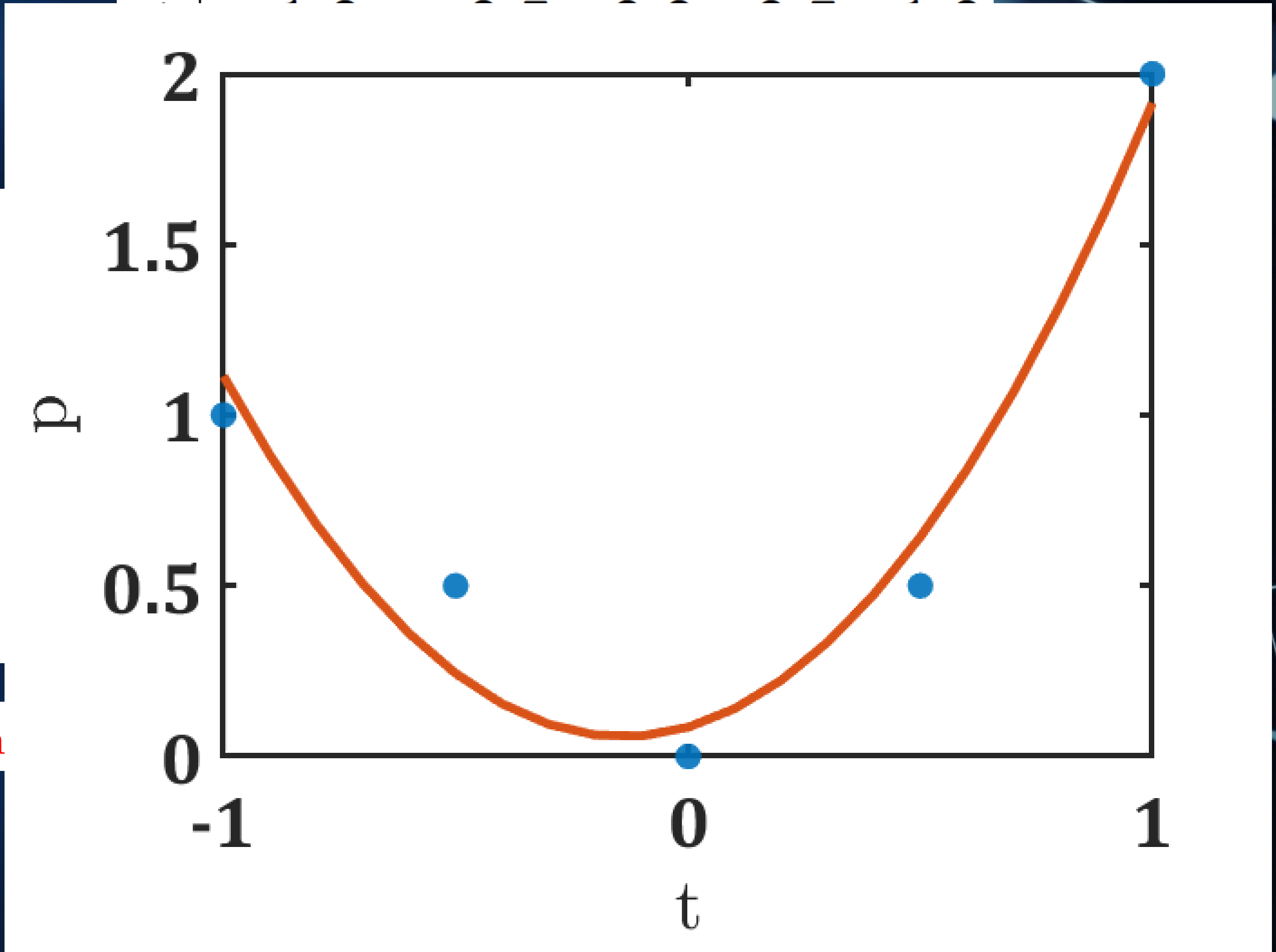
$$p(t) = 0.086 + 0.4t + 1.4t^2$$

Plot the data and the estimated curve

Solving LLS problems

$$2A^T b = 2A^T A x$$

Ax



LLS an

Data fitting and cool plot

What kind of transformation leaves least squares solution unchanged?

Square matrix Q is *orthogonal* if $Q^T Q = I$

Multiplying both sides of least squares problem by orthogonal matrix does not change solution.

Square matrix Q is *orthogonal* if $Q^T Q = I$

Preserves Euclidean norm, since

$$\|Qv\|_2^2 = (Qv)^T Qv = v^T Q^T Qv = v^T v = \|v\|_2^2$$

Multiplying both sides of least squares problem by orthogonal matrix does not change solution.

Square matrix Q is *orthogonal* if $Q^T Q = I$

Preserves Euclidean norm, since

$$\|Qv\|_2^2 = (Qv)^T Qv = v^T Q^T Qv = v^T v = \|v\|_2^2$$

Target: To transform general least squares problem to triangular form using orthogonal transformation.

QR factorization

QR factorization

Given $m \times n$ matrix $\boldsymbol{A}$, with $m > n$, we seek $m \times m$ orthogonal matrix $\boldsymbol{Q}$ such that

$$\boldsymbol{A} = \boldsymbol{Q} \begin{bmatrix} \boldsymbol{R} \\ \boldsymbol{O} \end{bmatrix},$$

with $\boldsymbol{R}$ $n \times n$ and upper triangular

QR factorization

Linear least squares problem $\boldsymbol{Ax} \cong \boldsymbol{b}$ transformed into triangular least squares problem

$$Q^T \boldsymbol{A}\boldsymbol{x} = \begin{bmatrix} \boldsymbol{R} \\ \boldsymbol{O} \end{bmatrix} \boldsymbol{x} \cong \begin{bmatrix} \boldsymbol{c}_1 \\ \boldsymbol{c}_2 \end{bmatrix} = Q^T \boldsymbol{b},$$

$A = Q\,[R\ O]^T$

$Q^T Q = I$

QR factorization

Linear least squares problem $Ax \cong b$ transformed into triangular least squares problem

$$Q^T A x = \begin{bmatrix} R \\ O \end{bmatrix} x \cong \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = Q^T b,$$

$A = Q\,[R\ O]^T$
$Q^T Q = I$

which has same solution, since $\|r\|_2^2 =$

$$\begin{aligned} \|b - Ax\|_2^2 &= \|b - Q \begin{bmatrix} R \\ O \end{bmatrix} x\|_2^2 = \|Q^T b - \begin{bmatrix} R \\ O \end{bmatrix} x\|_2^2 \\ &= \|c_1 - Rx\|_2^2 + \|c_2\|_2^2 \end{aligned}$$

because orthogonal transformation preserves Euclidean norm

QR factorization

To compute QR factorization of $m \times n$ matrix $\boldsymbol{A}$, with $m > n$, annihilate subdiagonal entries of successive columns of $\boldsymbol{A}$, eventually reaching upper triangular form

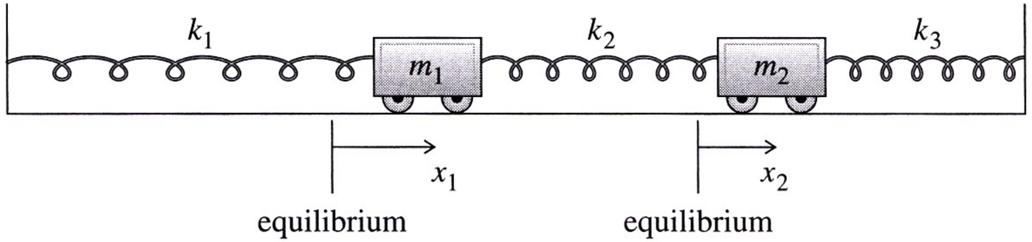
Similar to LU factorization by Gaussian elimination, but uses orthogonal transformations instead of elementary elimination matrices

Possible methods include

- Householder transformations
- Givens rotations
- Gram-Schmidt orthogonalization

Two Masses and Three Springs

- Here consider a system with two oscillators coupled together. The coupling can be strong or weak.
- We will limit the discussion to oscillators obeying Hooke's Law, and without friction. It is a special case, but one with a wide application.
- Consider the situation shown in the figure at right. There are two cars of masses m_1 and m_2, and three springs of spring constants k_1, k_2 and k_3, and we want to obtain the equations of motion for the two cars.



- We could use the Lagrangian formalism, but let's use the Newtonian approach first. The equilibrium positions of the two cars are shown by the lines, and we will use coordinates x_1 and x_2 relative to those.
- The forces on m_1 are k_1x_1 to the left, and $k_2(x_2 - x_1)$ to the right, so its equation of motion is

$$\begin{aligned} m_1\ddot{x}_1 &= -k_1x_1 + k_2(x_2 - x_1) \\ &= -(k_1 + k_2)x_1 + k_2x_2, \end{aligned}$$

Likewise:

$$m_2\ddot{x}_2 = k_2x_1 - (k_2 + k_3)x_2.$$

Two Masses and Three Springs

- The two **coupled** equations of motion:

$$m_1\ddot{x}_1 = -(k_1 + k_2)x_1 + k_2x_2$$

$$m_2\ddot{x}_2 = k_2x_1 - (k_2 + k_3)x_2,$$

can be written more compactly using matrix notation, as $\mathbf{M}\ddot{\mathbf{x}} = -\mathbf{K}\mathbf{x},$ where

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \mathbf{M} = \begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix}, \quad \text{and} \quad \mathbf{K} = \begin{bmatrix} k_1 + k_2 & -k_2 \\ -k_2 & k_2 + k_3 \end{bmatrix}.$$

Two Masses and Three Springs

- The two **coupled** equations of motion:

$$m_1\ddot{x}_1 = -(k_1 + k_2)x_1 + k_2 x_2$$
$$m_2\ddot{x}_2 = k_2 x_1 - (k_2 + k_3)x_2,$$

  can be written more compactly using matrix notation, as $\mathbf{M}\ddot{\mathbf{x}} = -\mathbf{K}\mathbf{x},$ where

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \qquad \mathbf{M} = \begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix}, \qquad \text{and} \qquad \mathbf{K} = \begin{bmatrix} k_1 + k_2 & -k_2 \\ -k_2 & k_2 + k_3 \end{bmatrix}.$$

- Notice that this is a generalization of the single oscillator, which you can see by setting k_2 and $k_3 = 0$. You then get a single equation. Note also that if the coupling spring, $k_2 = 0$, then the two equations become uncoupled and describe two separate oscillators.

Two Masses and Three Springs

- The two **coupled** equations of motion:

$$m_1\ddot{x}_1 = -(k_1 + k_2)x_1 + k_2 x_2$$
$$m_2\ddot{x}_2 = k_2 x_1 - (k_2 + k_3)x_2,$$

can be written more compactly using matrix notation, as $\mathbf{M}\ddot{\mathbf{x}} = -\mathbf{K}\mathbf{x}$, where

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \mathbf{M} = \begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix}, \quad \text{and} \quad \mathbf{K} = \begin{bmatrix} k_1 + k_2 & -k_2 \\ -k_2 & k_2 + k_3 \end{bmatrix}.$$

- Notice that this is a generalization of the single oscillator, which you can see by setting k_2 and $k_3 = 0$. You then get a single equation. Note also that if the coupling spring, $k_2 = 0$, then the two equations become uncoupled and describe two separate oscillators.
- We will find complex solutions $z(t) = ae^{i\omega t}$, but you can imagine that we might have more than one frequency of oscillation, since we have two ms and 3 ks.
- It turns out that we only need to assume one frequency $\omega = \sqrt{k/m}$, initially,

but we will arrive at an equation for ω that is satisfied by more than one frequency.

- Let's try the solutions: $$\mathbf{z}(t) = \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} e^{i\omega t} = \mathbf{a}e^{i\omega t}, \quad \text{where} \quad \mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \alpha_1 e^{-i\delta_1} \\ \alpha_2 e^{-i\delta_2} \end{bmatrix}.$$