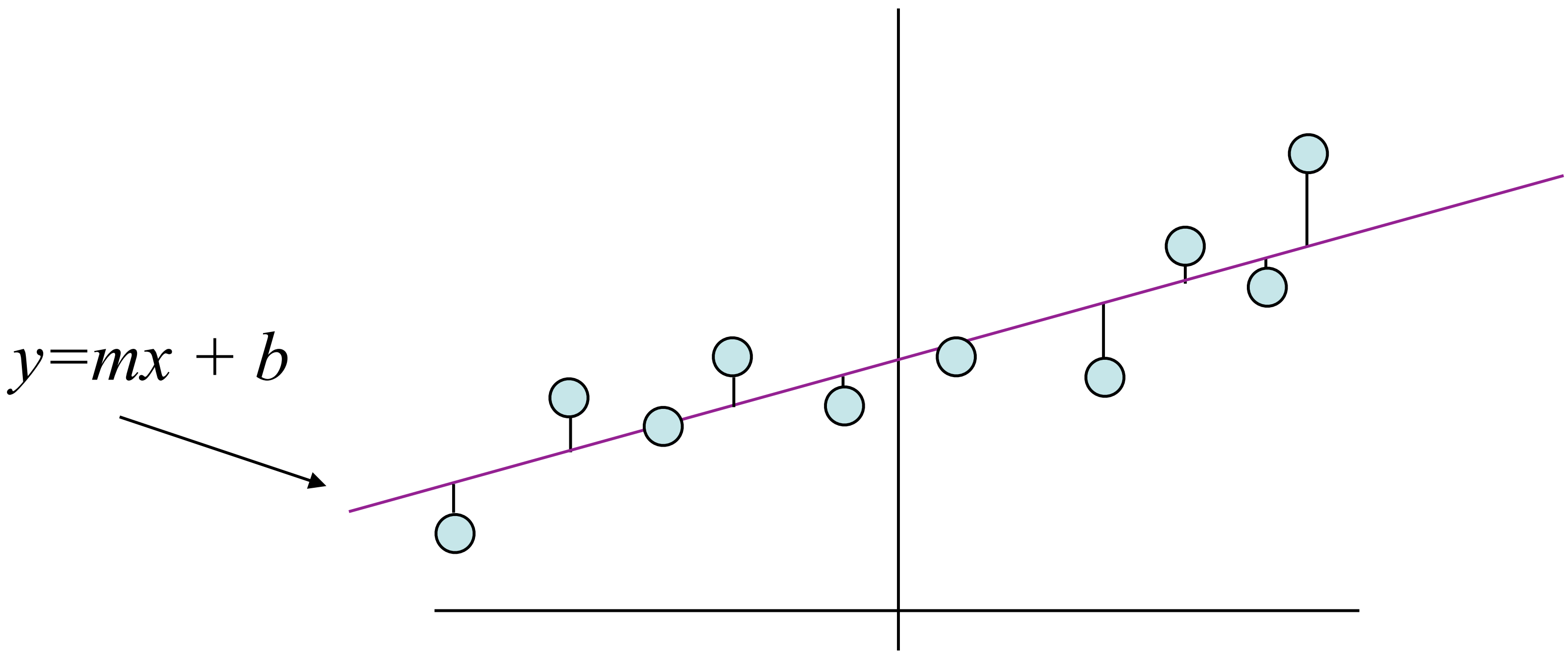
CS 477/577

Lecture Four

Homogenous least squares

Review

Non-homogeneous linear least squares (example two---naïve line fitting)



The error contribution for each point is the vertical distance (illustrated) squared.

We fit the line by tweaking m and b so that the sum of the contributions is as small as possible.

Review

Line fitting using non-homogeneous linear least squares

- We are minimizing $$\sum_i \left(y_i - \left(m x_i + b \right) \right)^2$$

- Only the vertical deviations count

- Asymmetric with respect to the axis
  - If you switch x and y, you will get a different answer

- Terminology note: This is standard *linear regression*

- Modeling note: The assumption here is that y_i are generated from from x_i, m, b, with added Gaussian noise.

Homogeneous linear least squares

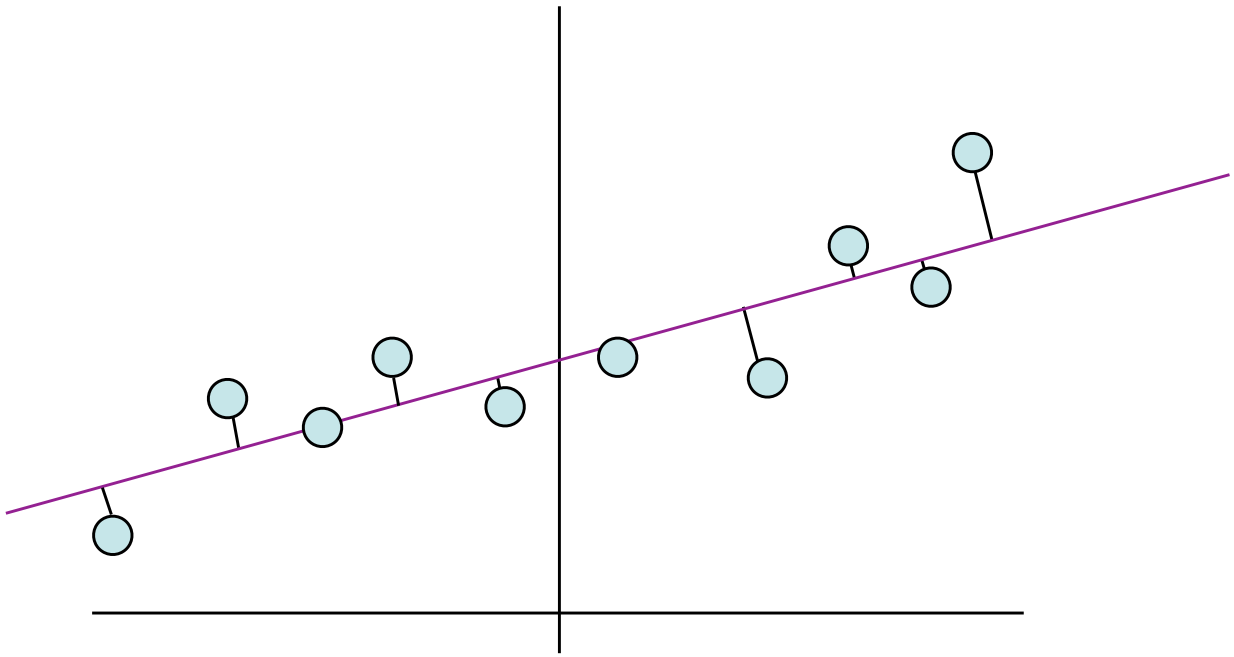
- A more symmetric alternative for line fitting is homogeneous least squares

- We need it for geometric camera calibration, but we will do line fitting first so we can use it in HW3

- For line fitting, the error is will now be the **perpendicular** distance from the data point to line hypothesis

- Terminology note: This is sometimes called *total least squares* (TLS)

Homogeneous linear least squares



Line fitting with homogenous linear least squares

A symmetric representation of a line is to draw a perpendicular from the origin to the line.

There are two variables:

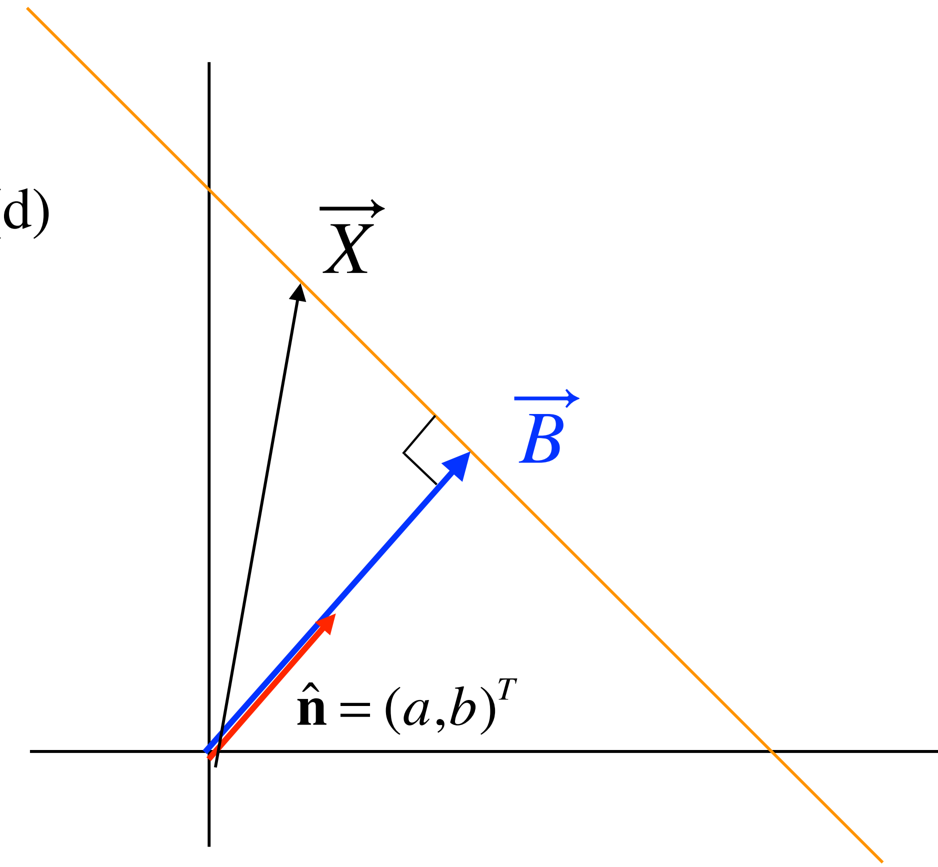
1) The direction of the perpendicular ($\hat{\mathbf{n}}$)
2) The distance from the origin to the line (d)

For a point $\overrightarrow{X} = (x, y)$ on the line

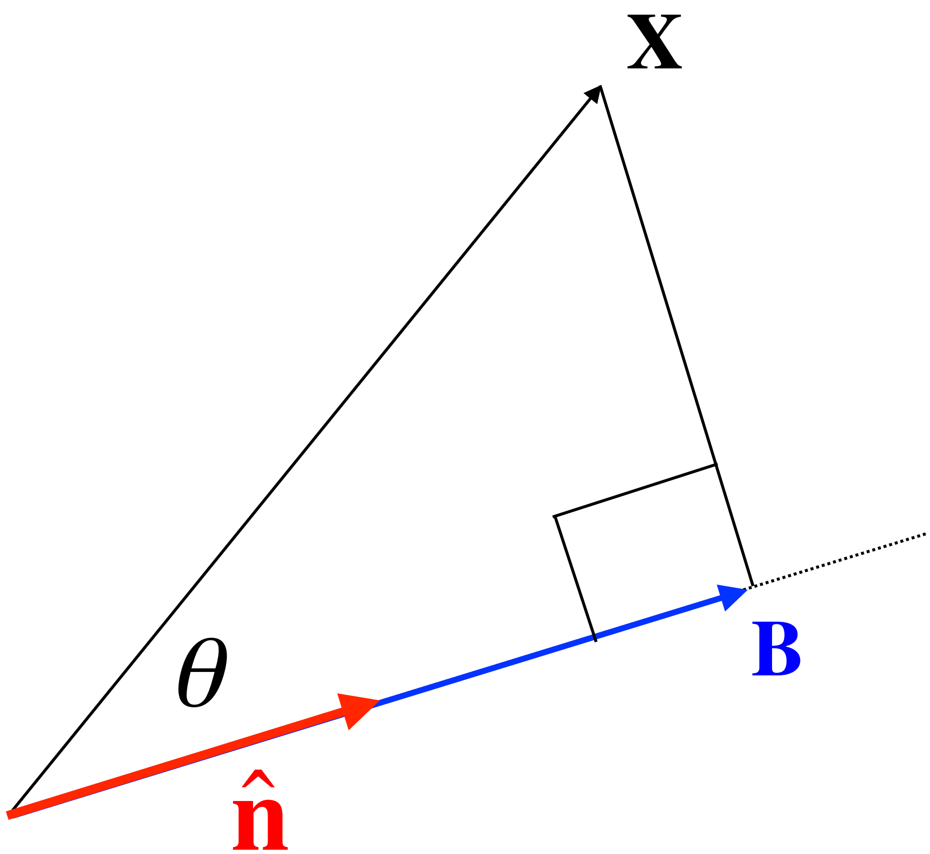
$\hat{\mathbf{n}} \cdot (x, y)$=d, where $d = |\overrightarrow{B}|$

Equivalently,

$$ax + by = d, \text{ where } a^2 + b^2 = 1$$



Dot product geometry



$\vec{B}$ is the perpendicular projection of $\vec{X}$ onto the line extending the unit vector $\mathbf{\hat{n}}$.

We derive that $\vec{X} \bullet \mathbf{\hat{n}} = \left|\vec{B}\right|$

From trigonometry $\left|\vec{B}\right| = (\cos\theta)\left|\vec{X}\right|$

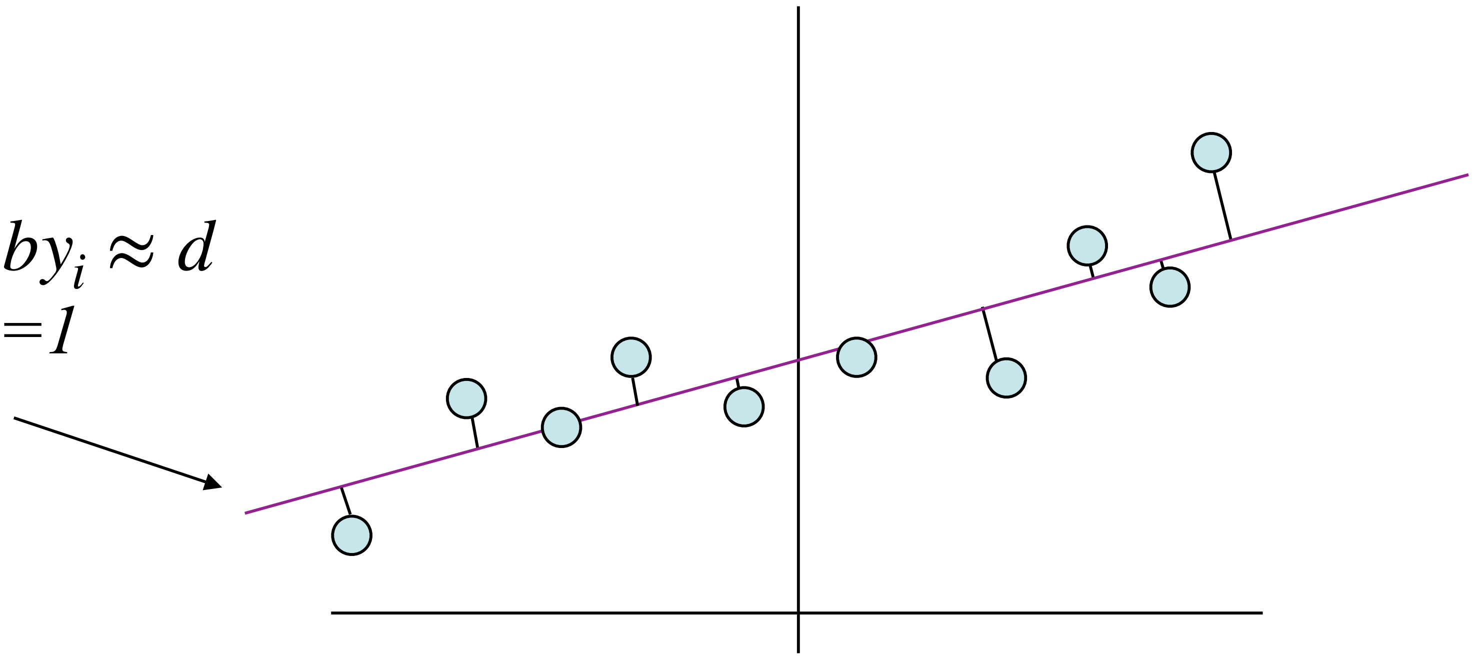
From the dot product $\vec{X} \bullet \vec{B} = (\cos\theta)\left|\vec{X}\right|\left|\vec{B}\right|$

and $\vec{X} \bullet \mathbf{\hat{n}} = (\cos\theta)\left|\vec{X}\right| = \left|\vec{B}\right|$ (divide by $\left|\vec{B}\right|$)

Important

Homogeneous linear least squares

$$ax_i + by_i \approx d$$
$$a^2+b^2=1$$



The error contribution for each point is the perpendicular distance (illustrated) squared.

We fit the line by tweaking a, b, and d so that the sum of the contributions is as small as possible subject to $a^2+b^2=1$.

Line fitting with homogenous linear least squares

Nice (symmetric) representation of a line:

$$ax + by = d, \text{ where } a^2 + b^2 = 1$$

So, for (x_i, y_i) on the line,

$$0 = |ax_i + by_i - d|$$

Key observation: The perpendicular distance from a point to that line is:

$$d_i = |ax_i + by_i - d|$$

(See upcoming slides for geometry)

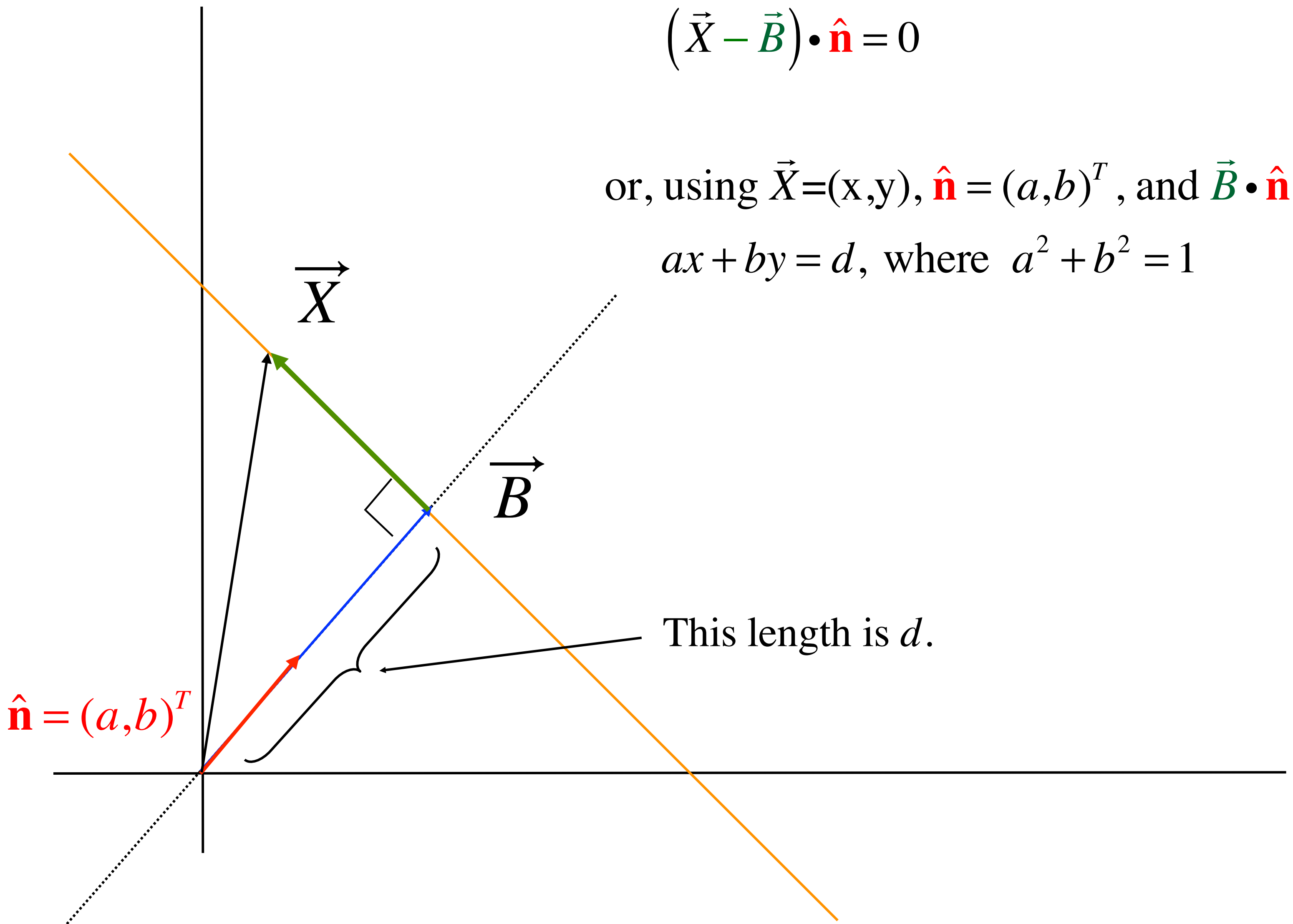
Line fitting with homogenous linear least squares

Equation for a point **x** on the line is:

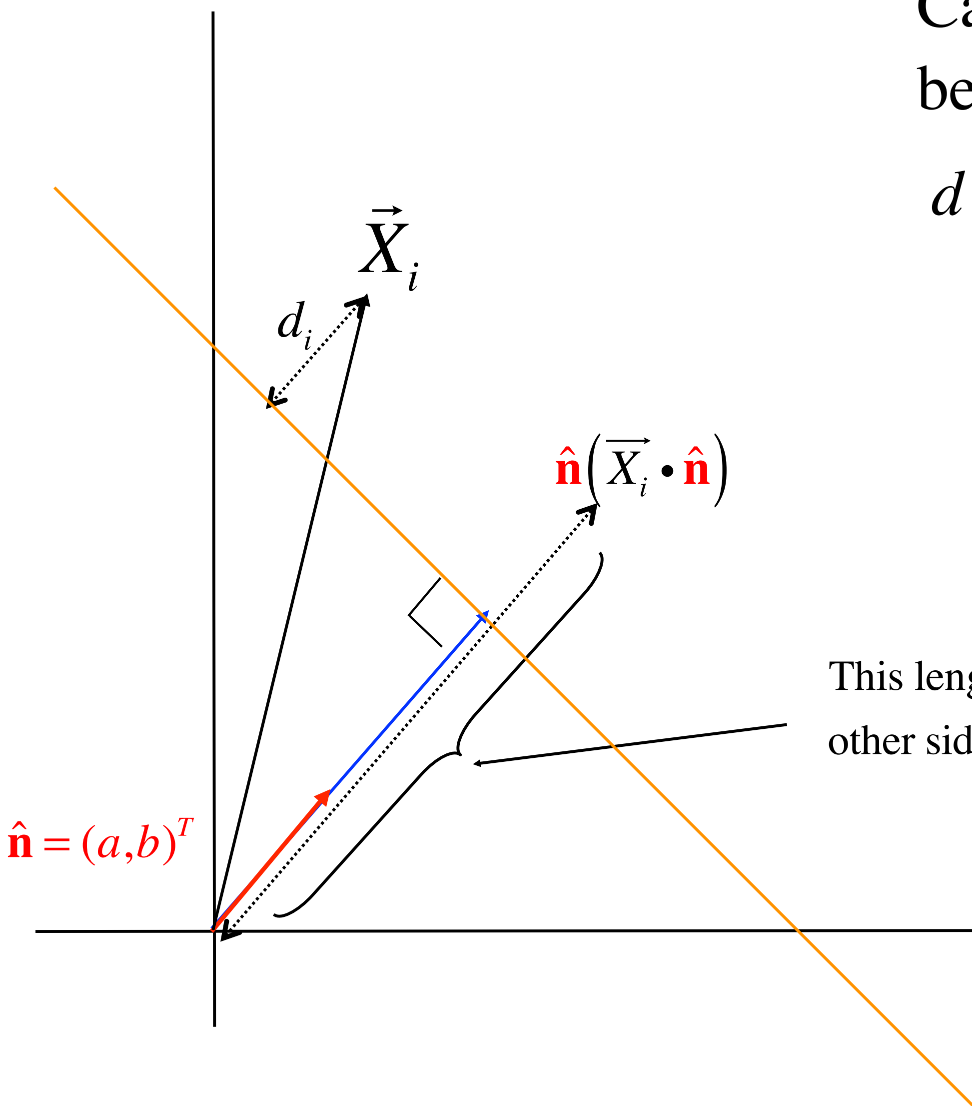
$$\left(\vec{X} - \vec{B}\right) \bullet \hat{\mathbf{n}} = 0$$

or, using $\vec{X}$=(x,y), $\hat{\mathbf{n}} = (a,b)^T$, and $\vec{B} \bullet \hat{\mathbf{n}} = d$,

$$ax + by = d, \text{ where } a^2 + b^2 = 1$$



Line fitting with homogenous linear least squares



Case where line goes between $\mathbf{x}_i$ and the origin

$$d + d_i = \vec{X}_i \bullet \hat{\mathbf{n}}$$

This length is $d_i + d$. If the point was on the other side of the line, we would have $-d_i + d$

Line fitting with homogenous linear least squares

Case where line goes between $\mathbf{x}_i$ and the origin

$$d + d_i = \vec{X}_i \bullet \hat{\mathbf{n}}$$

or

$$+d_i = \vec{X} \bullet \hat{\mathbf{n}} - d$$

Case where $\mathbf{x}_i$ is between the line and the origin

$$d - d_i = \vec{X} \bullet \hat{\mathbf{n}}$$

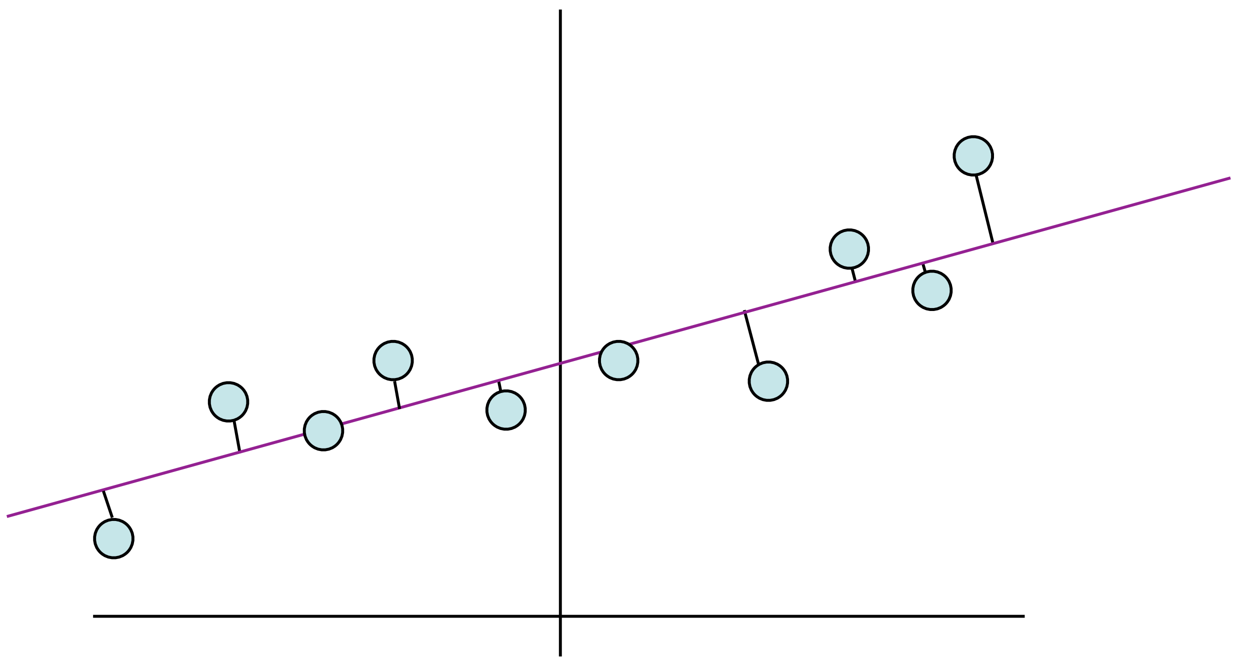
or

$$-d_i = \vec{X} \bullet \hat{\mathbf{n}} - d$$

$$|d_i| = |\vec{X}_i \bullet \hat{\mathbf{n}} - d|$$

Line fitting with homogenous linear least squares

$$E = \sum d_i^2 = \sum \left(d - ax_i - by_i\right)^2$$



Line fitting with homogenous linear least squares

$$E = \sum d_i^2 = \sum \left(d - a x_i - b y_i \right)^2$$

$$\frac{\partial E}{\partial d} = 2 \sum \left(d - a x_i - b y_i \right) = 0$$

So, $d = a\overline{x} + b\overline{y}$

Note that this means that the average point is on the line

Line fitting with homogenous linear least squares

$$E = \sum d_i^2$$

$$= \sum \left(d - ax_i - by_i\right)^2$$

$$= \sum \left(ax_i + by_i - d\right)^2$$ (reverse for future clarity)

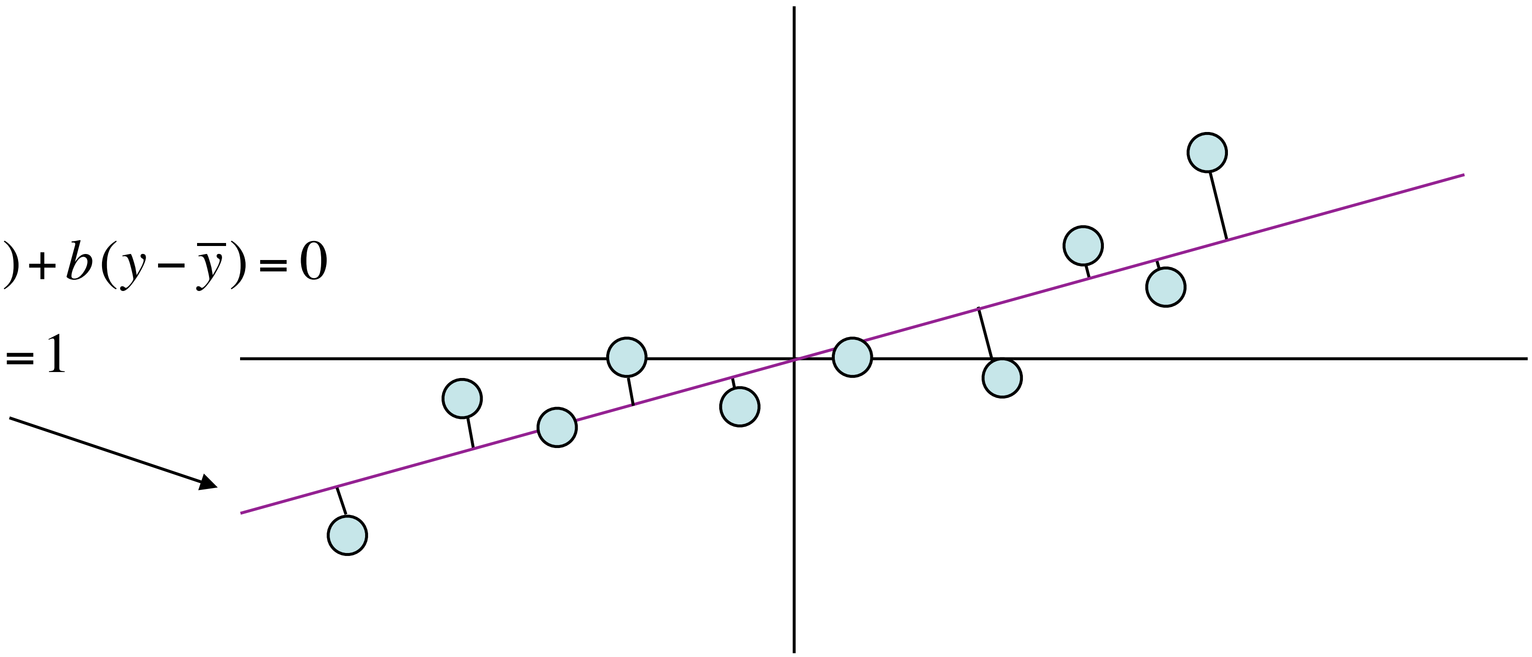
$$= \sum \left(ax_i - a\overline{x} + by_i - b\overline{y}\right)^2$$ $(d = a\overline{x} + b\overline{y})$

$$= \sum \left(\left(x_i - \overline{x}, y_i - \overline{y}\right) \bullet (a,b)\right)^2$$

Homogeneous linear least squares (new origin)

$$a(x-\overline{x})+b(y-\overline{y})=0$$

$$a^2+b^2=1$$



The error contribution for each point is the perpendicular distance (illustrated) squared.

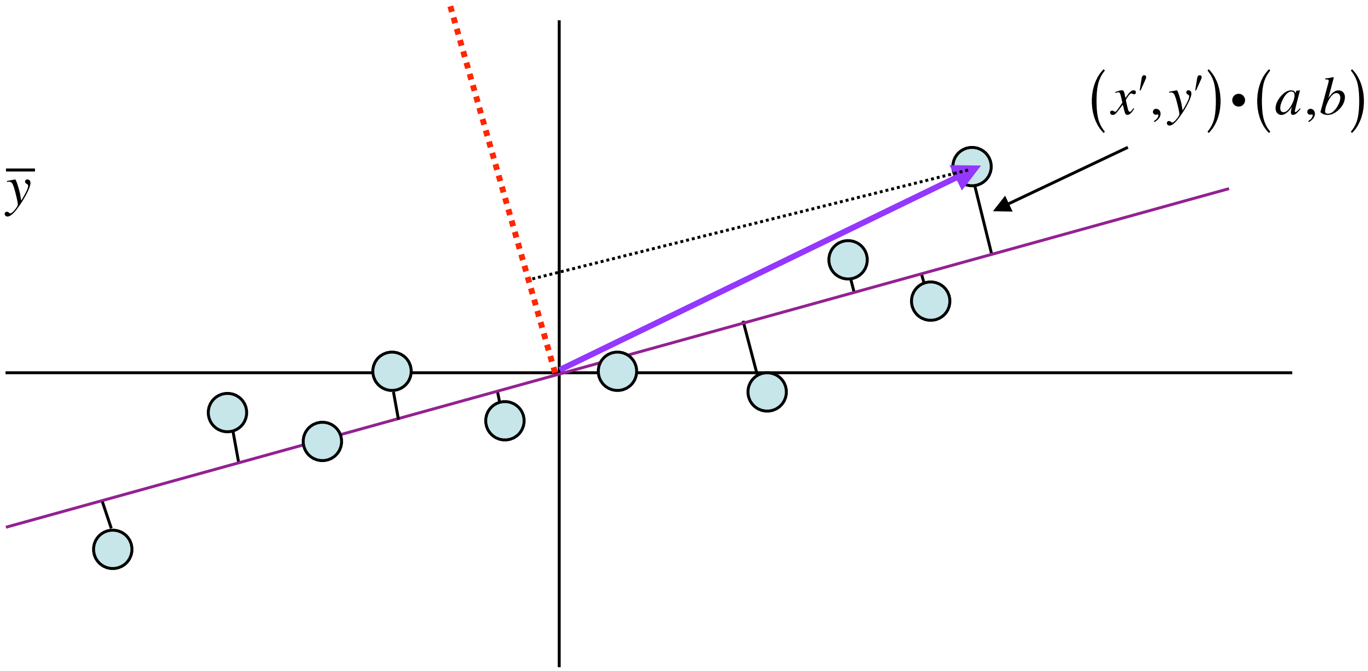
We fit the line by tweaking *a*, *b*, and *d* so that the sum of the contributions is as small as possible subject to $a^2+b^2=1$.

Homogeneous linear least squares (new origin)

$$x' = x - \overline{x}, \quad y' = y - \overline{y}$$

$$ax' + by' = 0$$

$$a^2 + b^2 = 1$$



The error contribution for each point is the perpendicular distance (illustrated) squared.

We fit the line by tweaking *a*, *b*, and *d* so that the sum of the contributions is as small as possible subject to $a^2+b^2=1$.

Line fitting with homogenous linear least squares

$$E = \sum \left(\left(x_i - \overline{x}, y_i - \overline{y} \right) \bullet (a,b) \right)^2$$

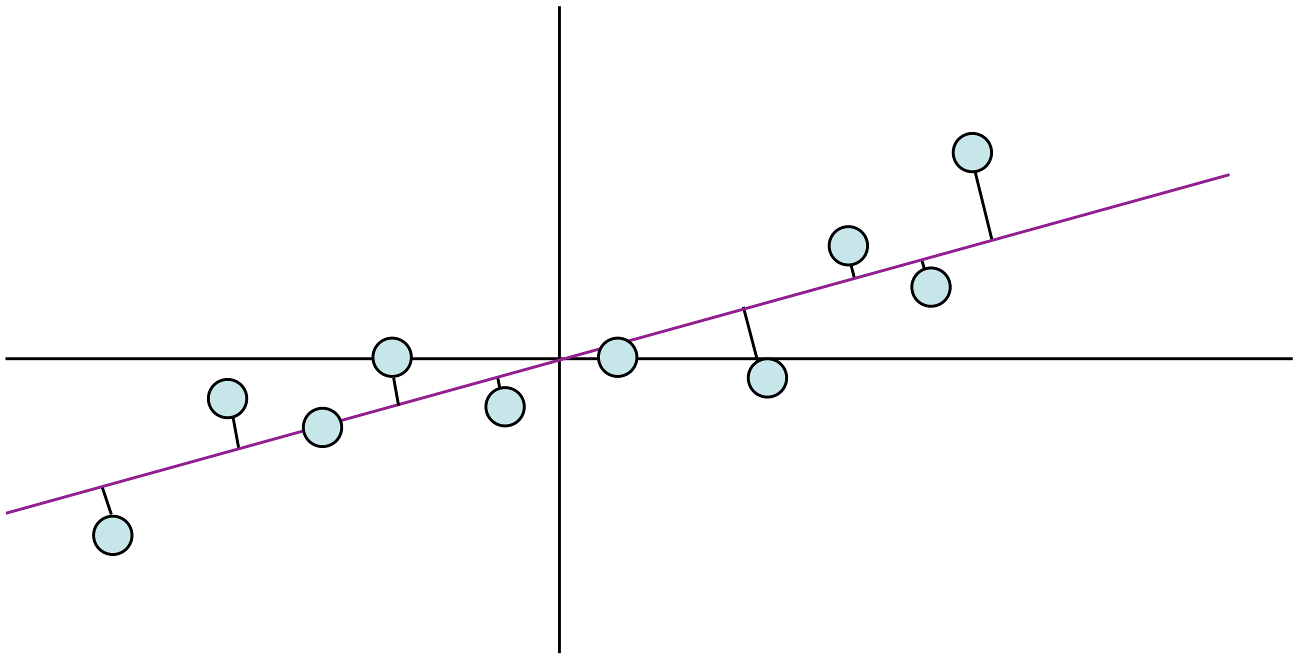
$$= \left| U\mathbf{n} \right|^2 ,$$

where $\quad U = \begin{pmatrix} x_1 - \overline{x} & y_1 - \overline{y} \\ \ldots & \ldots \\ x_n - \overline{x} & y_n - \overline{y} \end{pmatrix}$

and $\quad \mathbf{n} = \begin{pmatrix} a \\ b \end{pmatrix}$

So, we solve $U\mathbf{n}=0$ in the least squares sense, with the constraint that $\left| \mathbf{n} \right| = a^2 + b^2 = 1.$ (IE, use homogenous least squares)

Homogeneous linear least squares (new origin)



Does this look familiar (especially if you are a grad student)?

Math aside

Homogenous linear least squares

$$\text{Solve} \quad U\mathbf{x} \cong \mathbf{0} \quad \text{subject to} \quad \|\mathbf{x}\| = 1$$

Again, **there is usually no exact solution**.

Least squares solution is the value of **x** so that the magnitude of $U\mathbf{x}$ is as close zero as possible.

Math aside

Homogenous linear least squares

The least squares problem is thus

$$\text{Minimize } \|U\mathbf{x}\| \text{ subject to } \|\mathbf{x}\| = 1$$

This is solved by **magic** (but I can try to demystify this a little)

Wisdom from tea dipper handle

Math aside

Homogenous linear least squares

The least squares problem is thus

$$\text{Minimize } \|U\mathbf{x}\| \text{ subject to } \|\mathbf{x}\| = 1$$

This is solved by **magic** (but I will can try to demystify this a little)

Important

Specifically, the minimum is reached when $\mathbf{x}$ is set to the eigenvector corresponding to the minimum eigenvalue of $U^{\mathrm{T}}U$.

Important

Pragmatic solution of homogenous systems

To solve $U\mathbf{x} \cong \mathbf{0}$ subject to $\left\|\mathbf{x}\right\| = 1$

In Matlab, form $Y = U^{T} U$

Then use `eig`() to get the eigenvalues and eigenvectors of Y.

$\mathbf{x}$ is the eigenvector corresponding to the smallest eigenvalue. In Matlab, the smallest eigenvalue is usually first, but it pays to check.

("proof" --- try 1,000,000 random solutions to see if you can do better)

Optional material.

Justifying homogenous linear least squares (précis)

Solve $U\mathbf{x} \cong \mathbf{0}$ subject to $\|\mathbf{x}\| = 1$

IE, minimize $(U\mathbf{x})^T(U\mathbf{x}) = \mathbf{x}^T U^T U\mathbf{x}$ where $\mathbf{x}^T\mathbf{x} = 1$

$U^T U = VDV^T$ (Eigenvector decomposition)

D is diagonal, with postive real eigenvalues as elements

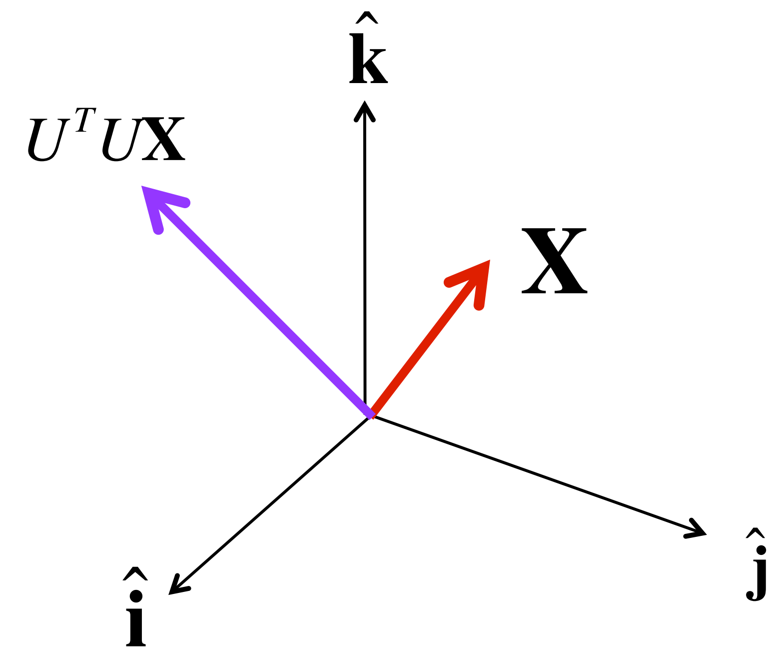
V and V^T are orthogonal with eigenvector columns/rows

Orthogonal transforms do not change the magnitude of vectors

Optional material.

Review of linear operators

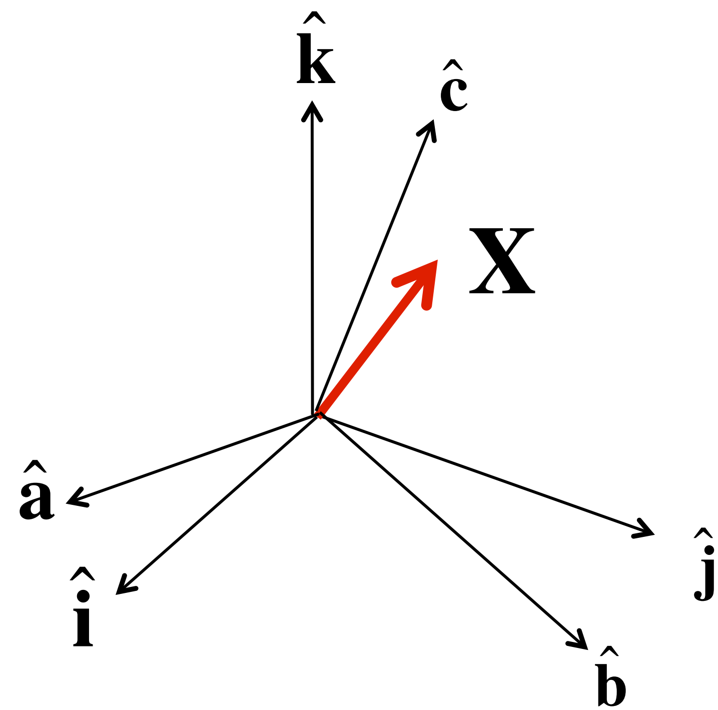
A matrix, e.g., A=U^TU, transforms a vector



Optional material.

Review of Change of Basis

A change of orthonormal basis only rewrites the coordinates, it does not change **X**, and in particular, does not change its length.



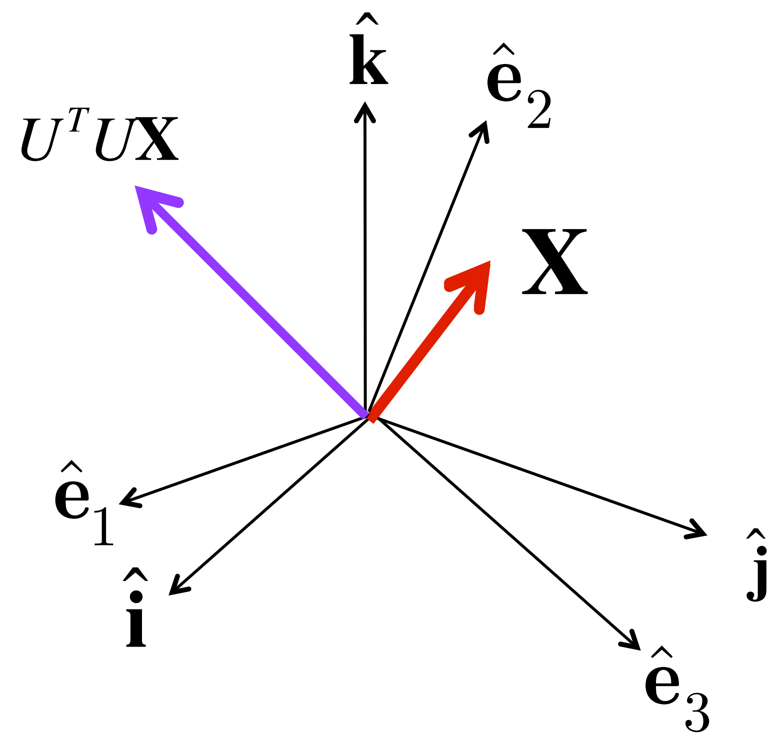
Optional material.

Using the Eigenvector Basis for $U^T U$

$U^T U = VDV^T$

(Eigenvector decomposition)

$$\left(U^T U\right)\mathbf{X} = \underbrace{V\left\{\underbrace{D\left(\underbrace{V^T\mathbf{X}}_{\text{rewrite coordinates in eigen-basis}}\right)}_{\text{stretch coordinates by eigenvalues (changes magnitude)}}\right\}}_{\text{rewrite coordinates as original coordinates}}$$

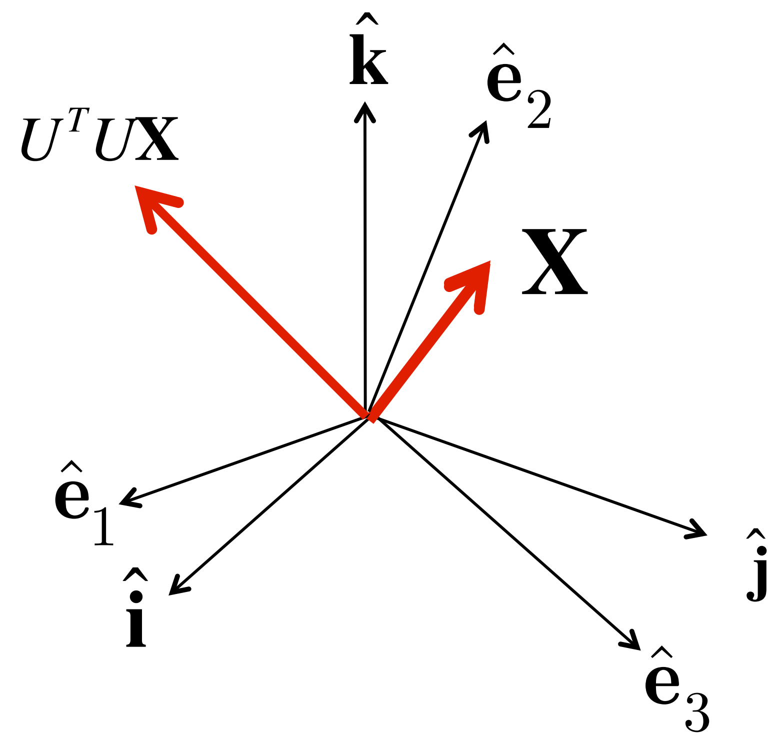


Using the Eigenvector Basis for U^TU

$U^T U = VDV^T$

(Eigenvector decomposition)

In this coordinate rewrite, we stretch the vector in each dimension according to D that has the eigenvalues for the entries.



Optional material.

Using the Eigenvector Basis for $U^T U$

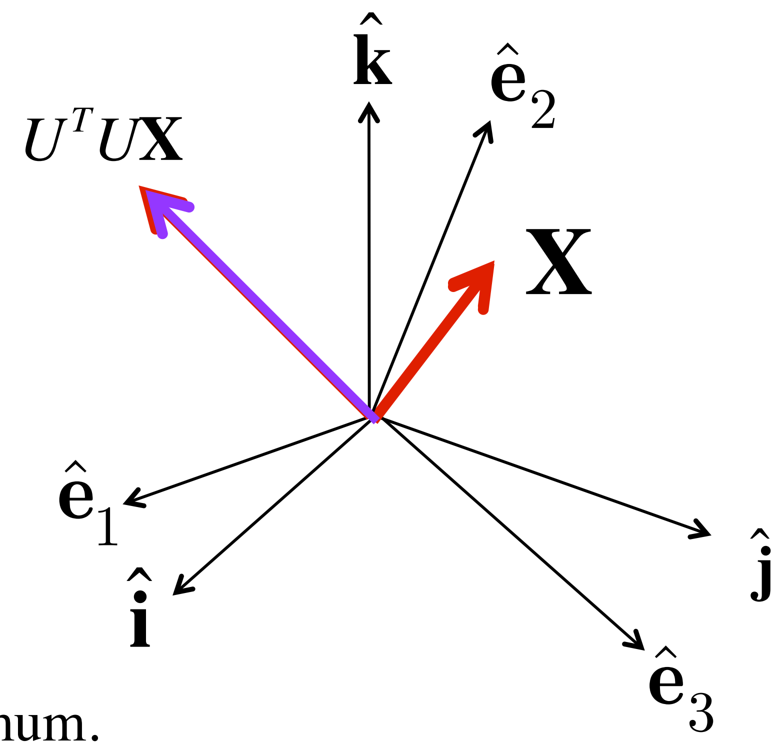
$U^T U = V D V^T$

(Eigenvector decomposition)

In this coordinate rewrite, we stretch the vector in each dimension according to D that has the eigenvalues for the entries.

So, assuming the first eigenvalue is the smallest, (1,0,0) in eigenvector coordinates, gives the minimum.

So $\mathbf{X}$ should be $\hat{\mathbf{e}}_{\min}$ in the original coordinates



Optional material.

Justifying homogenous linear least squares (précis)

Solve $U\mathbf{x} \cong \mathbf{0}$ subject to $\|\mathbf{x}\| = 1$

IE, minimize $(U\mathbf{x})^T (U\mathbf{x}) = \mathbf{x}^T U^T U\mathbf{x}$ where $\mathbf{x}^T\mathbf{x} = 1$

$U^T U = VDV^T$ (Eigenvector decomposition)

D is diagonal, with postive real eigenvalues as elements

V and V^T are orthogonal with eigenvector columns/rows

Orthogonal transforms do not change the magnitude of vectors

$$(U\mathbf{x})^T (U\mathbf{x}) = \left(V^T\mathbf{x}\right)^T D\left(V^T\mathbf{x}\right)$$

The effect in the eigenbasis is to stretch components by the (positive) values in D

The best you can do is to put all the weight on the smallest value of D .

This is achieved by the eigenvector corresponding to the minimal eigenvalue in D.

Optional material.

Justifying homogenous linear least squares (précis)

$$(U\mathbf{x})^T (U\mathbf{x}) = \left(V^T\mathbf{x}\right)^T D\left(V^T\mathbf{x}\right)$$ (from before)

$$= \left(V^T\mathbf{x}\right)^T \sqrt{D}\sqrt{D}\left(V^T\mathbf{x}\right)$$ (we *know* that the elements of D are positive)

$$= \left(\sqrt{D}V^T\mathbf{x}\right)\bullet\left(\sqrt{D}V^T\mathbf{x}\right)$$

$$= \left\|\sqrt{D}V^T\mathbf{x}\right\|^2$$

Additional supplementary material with further details justifying the formula for homogenous least squares (not covered in the asynchronous pre-recorded lecture).

Justifying homogenous linear least squares (pedantic)

Because we solve $U\mathbf{x}=\mathbf{0}$ as best we can, the error vector is $U\mathbf{x}$

The squared error is then

$$\left(U\mathbf{x}\right)^{\mathrm{T}}\left(U\mathbf{x}\right)=\mathbf{x}^{\mathrm{T}}\left(U^{\mathrm{T}}U\right)\mathbf{x}$$

Supplementary material

Justifying homogenous linear least squares (pedantic)

$$\text{Since } U^T U \text{ is positive semidefinite, the eigenvalues are } \mathbf{non-negative}$$

Recall that a matrix A is positive semidefinite if $\mathbf{x}^T A \mathbf{x}$ is never negative.

(Clearly $U^T U$ it is positive semidefinite because $\mathbf{x}^T U^T U \mathbf{x}$ is $|U\mathbf{x}|^2$

Supplementary material

Further technical comments

If the model is good, then U will **approximate** a matrix of deficient column rank because there should exist a non-zero $\mathbf{x}$ that solves U$\mathbf{x}$=$\mathbf{0}$.

We force the solution process to embody the assumption that the fact that U^TU *appears* to be of full rank is due to measurement error. This assumption helps separate the part of U that is due to errors from the part that is due to the model.

This why we say that U^TU is **semi**-positive definite, and *not* positive definite. We assume that there may a solution to Ux=0, that is distinctly non-zero .

(A matrix, A, is positive definite if, $\mathbf{x}^T A \mathbf{x}$ is never negative, **and** $\mathbf{x}^T A \mathbf{x} = \mathbf{0}$ means that $\mathbf{x}=\mathbf{0}$.)

Supplementary material

Justifying homogenous linear least squares (pedantic)

Since $U^T U$ it is positive semidefinite, the eigenvalues are **non – negative**

Recall that a matrix A has an eigenvector, $\mathbf{e}$, with eigenvalue λ if

$$A\mathbf{e}=\lambda\mathbf{e}$$

Diagonalization:

$U^{\mathbf{T}}U = V\Lambda V^T$ where V is an orthonormal basis made of the eigenvectors, $\mathbf{e}_i$, of $U^{\mathbf{T}}U$, and Λ is a diagonal matrix of the eigenvalues

Supplementary material

Justifying homogenous linear least squares (pedantic)

Since $U^T U$ it is positive semidefinite, the eigenvalues are **non – negative**

We will write them as $\lambda_i = \sigma_i^2$.

Note: The book (at least my copy) uses λ_i^2 in the equation at the top of page 41 which is confusing. The coefficients, normally denoted, λ_i are in fact equal to the **square** of the "singular values of U", which usually are denoted by σ_i

Justifying homogenous linear least squares (pedantic)

We can write **x** in terms of the eigenvector basis:

$$\mathbf{x}=\sum u_{\mathrm{i}}\mathbf{e}_{\mathrm{i}} \quad \text{where} \quad \sum u_i^2 = 1 \qquad \text{(why?)}$$

Supplementary material

Justifying homogenous linear least squares (pedantic)

We can write **x** in terms of the eigenvector basis:

$$\mathbf{x}=\sum u_{\mathrm{i}}\mathbf{e}_{\mathrm{i}} \quad \text{where} \quad \sum u_i^2 = 1 \qquad \text{(why?)}$$

$$\text{(Because } \mathbf{x}^{\mathrm{T}}\mathbf{x} = \sum u_{\mathrm{j}}\mathbf{e}_j^{\mathrm{T}} \sum u_{\mathrm{i}}\mathbf{e}_i = \sum \sum u_{\mathrm{i}}u_{\mathrm{j}}\mathbf{e}_j^{\mathrm{T}}\mathbf{e}_i = \sum u_i^2 \text{)}$$

$$\text{(and } |\mathbf{x}| = 1\text{)}$$

(Alternatively, appeal to the fact that the new basis is orthonormal, and so the length of **x** does not change).

Supplementary material

Justifying homogenous linear least squares (pedantic)

The error to minimize is $\mathbf{x}^{\mathbf{T}}\left(VDV^{T}\right)\mathbf{x}=\left(\mathbf{x}^{\mathbf{T}}V\right)D\left(V^{T}\mathbf{x}\right)$

Recall that $\mathbf{x}=\sum u_{\mathrm{i}}\mathbf{e}_{\mathrm{i}}$

And that the columns of V are the eigenvectors $\mathbf{e}_{i}$

So the elements of $V^{T}\mathbf{x}$ are ?

Supplementary material

Justifying homogenous linear least squares (pedantic)

The error to minimize is $\mathbf{x}^{\mathbf{T}}\left(VDV^{T}\right)\mathbf{x}=\left(\mathbf{x}^{\mathbf{T}}V\right)D\left(V^{T}\mathbf{x}\right)$

Recall that $\mathbf{x}=\sum u_{\mathrm{i}}\mathbf{e}_{\mathrm{i}}$

And that the columns of V are the eigenvectors $\mathbf{e}_{i}$

So the elements of $V^{T}\mathbf{x}$ are u_{i}

Justifying homogenous linear least squares (pedantic)

The error to minimize is $\mathbf{x}^{\mathbf{T}}\left(VDV^{T}\right)\mathbf{x}=\left(\mathbf{x}^{\mathbf{T}}V\right)D\left(V^{T}\mathbf{x}\right)$

And we have noted that the elements of $V^{T}\mathbf{x}$ are u_i

So the error is $?$

Supplementary material

Justifying homogenous linear least squares (pedantic)

The error to minimize is $\mathbf{x}^{\mathbf{T}}\left(VDV^{T}\right)\mathbf{x}=\left(\mathbf{x}^{\mathbf{T}}V\right)D\left(V^{T}\mathbf{x}\right)$

And we have noted that the elements of $V^{T}\mathbf{x}$ are u_i

So the error is $\sum u_i^2\lambda_i=\sum u_i^2\sigma_i^2$

Supplementary material

Justifying homogenous linear least squares (pedantic)

From the previous slide the error to be minized is $\sum u_i^2 \sigma_i^2$

We are stuck with the values σ_i^2 and $\sum u_i^2 = 1$

To make the error small, what can we do?

Supplementary material

Justifying homogenous linear least squares (pedantic)

From the previous slide the error to be minized is $\sum u_i^2\sigma_i^2$

We are stuck with the values σ_i^2 and $\sum u_i^2 = 1$

The best we can do is to set $u_i = 1$

for the minimum value of $\lambda_i = \sigma_i^2$ and zero for the others.