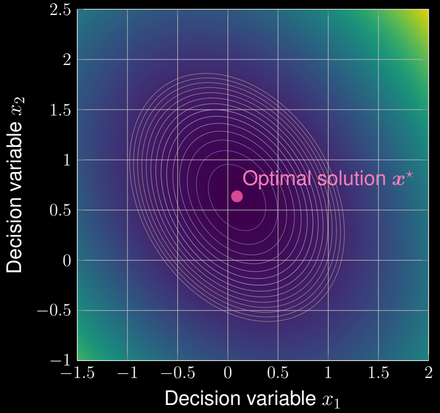
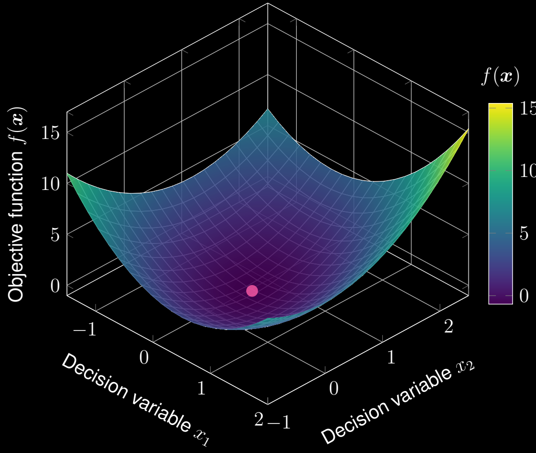
Quadratic Optimization

$$\min_{\boldsymbol{x}\in\mathbb{R}^2} \quad \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}^\top \boldsymbol{x} \qquad \text{with} \qquad \boldsymbol{A} = \begin{bmatrix} 4 & 1 \\ 1 & 3 \end{bmatrix} \quad \text{and} \quad \boldsymbol{b} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$



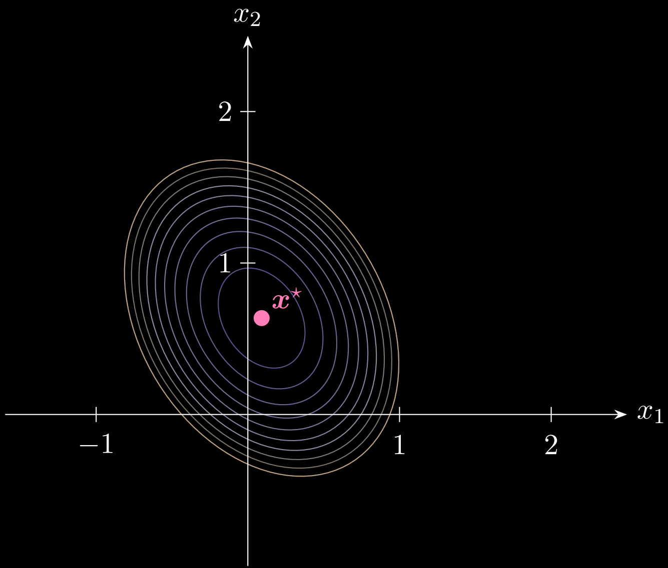


Quadratic Optimization

For any symmetric positive definite matrix $\boldsymbol{A}$ and vector $\boldsymbol{b}$:

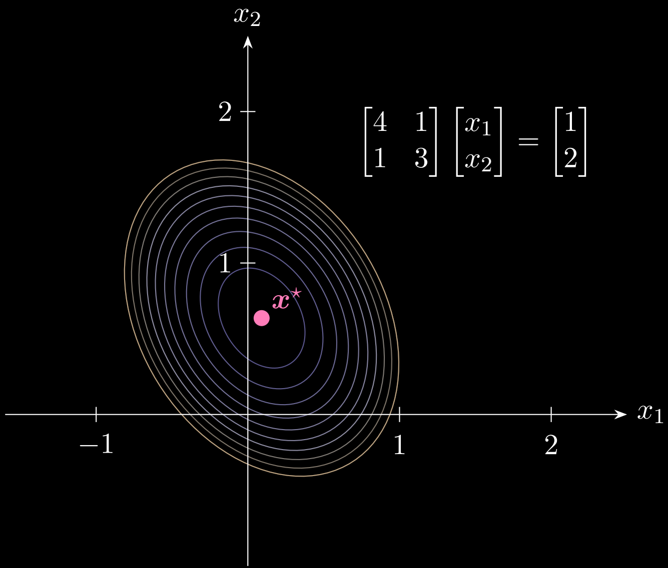
$$\min_{\boldsymbol{x}\in\mathbb{R}^2} \quad \underbrace{\frac{1}{2}\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}^\top \boldsymbol{x}}_{\text{Objective function } f} \quad \Rightarrow \quad \underbrace{\frac{\mathrm{d}f}{\mathrm{d}\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}}_{\text{First-order derivative}} = \boldsymbol{0} \quad \Rightarrow \quad \underbrace{\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}}_{\text{Linear system}}$$

Quadratic Optimization



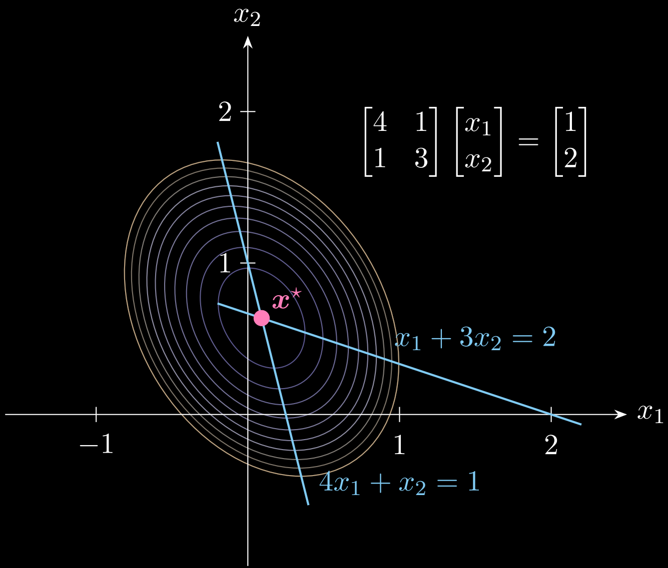
Equivalent Linear System $\boldsymbol{Ax} = \boldsymbol{b}$

$$\begin{bmatrix} 4 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$



Equivalent Linear System $\boldsymbol{Ax} = \boldsymbol{b}$

$$\begin{bmatrix} 4 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$



Update Equations

Essential idea: Conjugate gradient solves linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ iteratively.

❶ x is updated by

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \underbrace{\alpha_k}_{\text{step size (unknown)}} \boldsymbol{d}_k$$

with initial point $\boldsymbol{x}_0$.

❷ $\boldsymbol{d}_k$ is the search direction:

$$\boldsymbol{d}_{k+1} = \underbrace{\boldsymbol{r}_{k+1}}_{\text{residual}} + \underbrace{\beta_k}_{\text{coefficient (unknown)}} \boldsymbol{d}_k$$

❸ How to write down the update of residual $\boldsymbol{r}_{k+1}$?

$$\begin{aligned} \boldsymbol{x}_{k+1} &= \boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k \\ \Rightarrow \qquad \boldsymbol{A}\boldsymbol{x}_{k+1} &= \boldsymbol{A}\boldsymbol{x}_k + \alpha_k \boldsymbol{A}\boldsymbol{d}_k \\ \Rightarrow \qquad \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}_{k+1} &= \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}_k - \alpha_k \boldsymbol{A}\boldsymbol{d}_k \end{aligned}$$

❸ How to write down the update of residual $\boldsymbol{r}_{k+1}$?

$$
\begin{aligned}
& & \boldsymbol{x}_{k+1} &= \boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k \\
\Rightarrow & & \boldsymbol{A}\boldsymbol{x}_{k+1} &= \boldsymbol{A}\boldsymbol{x}_k + \alpha_k \boldsymbol{A}\boldsymbol{d}_k \\
\Rightarrow & & \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}_{k+1} &= \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}_k - \alpha_k \boldsymbol{A}\boldsymbol{d}_k \\
\Rightarrow & & \boldsymbol{r}_{k+1} &= \boldsymbol{r}_k - \alpha_k \boldsymbol{A}\boldsymbol{d}_k
\end{aligned}
$$

To summarize, we have the following update equations to solve $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k \qquad \text{(decision variables)}$$

$$\boldsymbol{r}_{k+1} = \boldsymbol{r}_k - \alpha_k \boldsymbol{A}\boldsymbol{d}_k \qquad \text{(residuals)}$$

$$\boldsymbol{d}_{k+1} = \boldsymbol{r}_{k+1} + \beta_k \boldsymbol{d}_k \qquad \text{(search direction)}$$

△ But how to write down the update equations of α_k and β_k?

Orthogonal residuals + Conjugate search directions ($\boldsymbol{A}$-orthogonality)

$$\boldsymbol{r}_{k+1}^\top \boldsymbol{r}_k = 0 \qquad \text{and} \qquad \boldsymbol{d}_{k+1}^\top \boldsymbol{A}\boldsymbol{d}_k = 0$$

for any iterations k and $k+1$.

❹ Orthogonal Residuals $\boldsymbol{r}_k^\top \boldsymbol{r}_{k+1} = 0$:

$$\boldsymbol{r}_k^\top \underbrace{(\boldsymbol{r}_k - \alpha_k \boldsymbol{A}\boldsymbol{d}_k)}_{=\boldsymbol{r}_{k+1}} = 0 \quad \Rightarrow \quad \alpha_k = \frac{\boldsymbol{r}_k^\top \boldsymbol{r}_k}{\boldsymbol{r}_k^\top \boldsymbol{A}\boldsymbol{d}_k}$$

❹ Orthogonal Residuals $\boldsymbol{r}_k^\top \boldsymbol{r}_{k+1} = 0$:

$$\boldsymbol{r}_k^\top \underbrace{(\boldsymbol{r}_k - \alpha_k \boldsymbol{A}\boldsymbol{d}_k)}_{=\boldsymbol{r}_{k+1}} = 0 \quad \Rightarrow \quad \alpha_k = \frac{\boldsymbol{r}_k^\top \boldsymbol{r}_k}{\boldsymbol{r}_k^\top \boldsymbol{A}\boldsymbol{d}_k}$$

- Denominator:

$$\underbrace{\boldsymbol{r}_k^\top \boldsymbol{A}\boldsymbol{d}_k = (\boldsymbol{d}_k - \beta_{k-1}\boldsymbol{d}_{k-1})^\top \boldsymbol{A}\boldsymbol{d}_k}_{\text{derived from the update equation of search direction}} = \boldsymbol{d}_k^\top \boldsymbol{A}\boldsymbol{d}_k - \beta_{k-1} \underbrace{\boldsymbol{d}_{k-1}^\top \boldsymbol{A}\boldsymbol{d}_k}_{=0 \text{ (conjugate)}} = \boldsymbol{d}_k^\top \boldsymbol{A}\boldsymbol{d}_k$$

- Update equation of step size α_k:

$$\alpha_k = \frac{\boldsymbol{r}_k^\top \boldsymbol{r}_k}{\boldsymbol{d}_k^\top \boldsymbol{A}\boldsymbol{d}_k}$$

❺ Conjugate search directions $\boldsymbol{d}_{k+1}^\top \boldsymbol{A}\boldsymbol{d}_k = 0$:

$$\underbrace{(\boldsymbol{r}_{k+1} + \beta_k \boldsymbol{d}_k)}_{=\boldsymbol{d}_{k+1}}{}^\top \boldsymbol{A}\boldsymbol{d}_k = 0 \quad \Rightarrow \quad \beta_k = -\frac{\boldsymbol{r}_{k+1}^\top \boldsymbol{A}\boldsymbol{d}_k}{\boldsymbol{d}_k^\top \boldsymbol{A}\boldsymbol{d}_k}$$

- Considering

$$\boldsymbol{A}\boldsymbol{d}_k = \frac{\boldsymbol{r}_k - \boldsymbol{r}_{k+1}}{\alpha_k} \qquad \alpha_k = \frac{\boldsymbol{r}_k^\top \boldsymbol{r}_k}{\boldsymbol{d}_k^\top \boldsymbol{A}\boldsymbol{d}_k}$$

- Update equation of coefficient $\boldsymbol{\beta}_k$:

$$\beta_k = \frac{\boldsymbol{r}_{k+1}^\top \boldsymbol{r}_{k+1}}{\boldsymbol{r}_k^\top \boldsymbol{r}_k}$$

Conjugate Gradient Method

To summarize, we have the following update equations to solve $\boldsymbol{Ax} = \boldsymbol{b}$

$$\alpha_k = \frac{\boldsymbol{r}_k^\top \boldsymbol{r}_k}{\boldsymbol{d}_k^\top \boldsymbol{A} \boldsymbol{d}_k} \qquad \text{(step size)}$$

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k \qquad \text{(decision variables)}$$

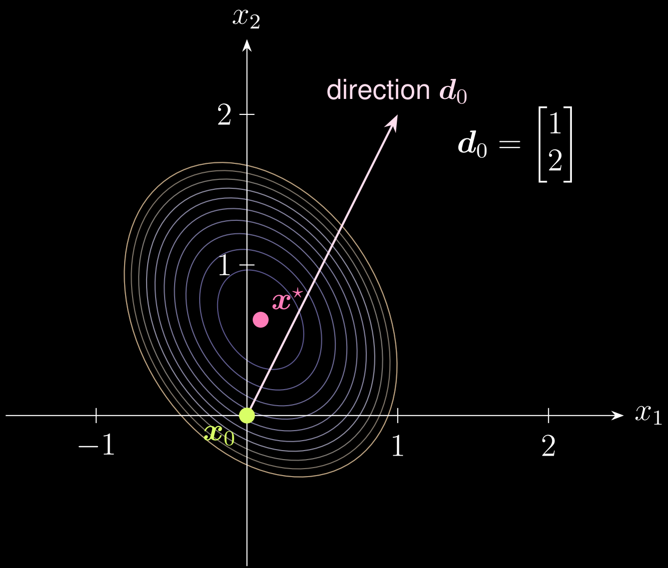
$$\boldsymbol{r}_{k+1} = \boldsymbol{r}_k - \alpha_k \boldsymbol{A} \boldsymbol{d}_k \qquad \text{(residuals)}$$

$$\beta_k = \frac{\boldsymbol{r}_{k+1}^\top \boldsymbol{r}_{k+1}}{\boldsymbol{r}_k^\top \boldsymbol{r}_k} \qquad \text{(coefficient)}$$

$$\boldsymbol{d}_{k+1} = \boldsymbol{r}_{k+1} + \beta_k \boldsymbol{d}_k \qquad \text{(search direction)}$$

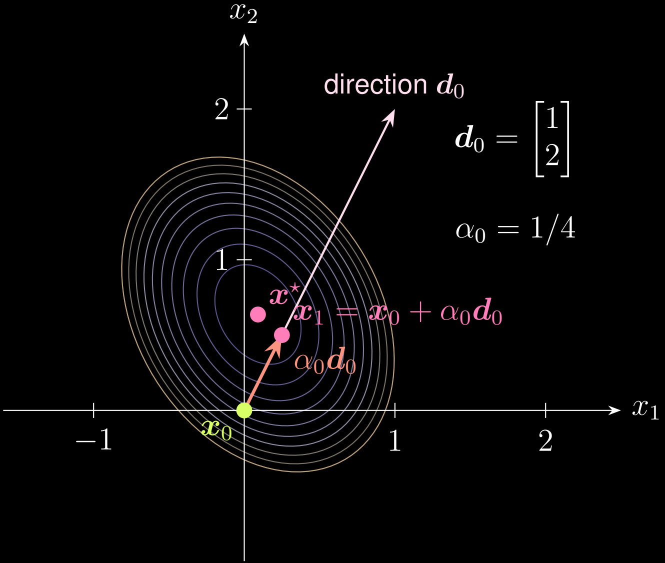
with $\boldsymbol{r}_0 = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}_0$ and $\boldsymbol{d}_0 = \boldsymbol{r}_0$ for initialization.

Conjugate Gradient Method



$$\boldsymbol{d}_0 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Conjugate Gradient Method

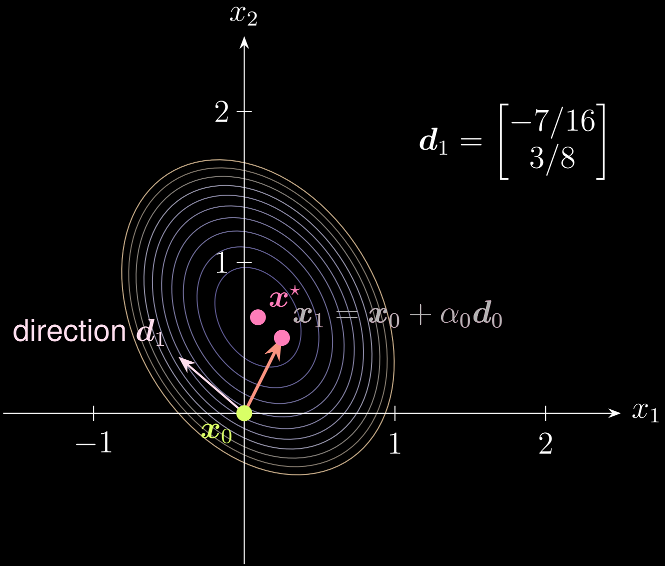


$$\boldsymbol{d}_0 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

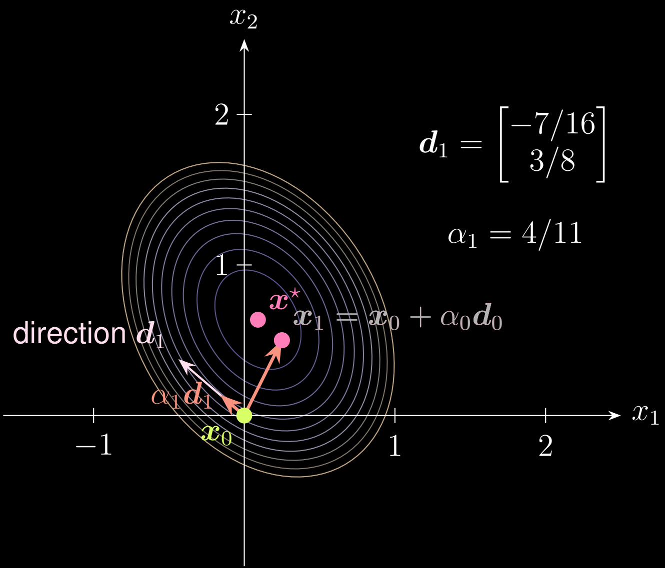
$$\alpha_0 = 1/4$$

Conjugate Gradient Method

$$\boldsymbol{d}_1 = \begin{bmatrix} -7/16 \\ 3/8 \end{bmatrix}$$



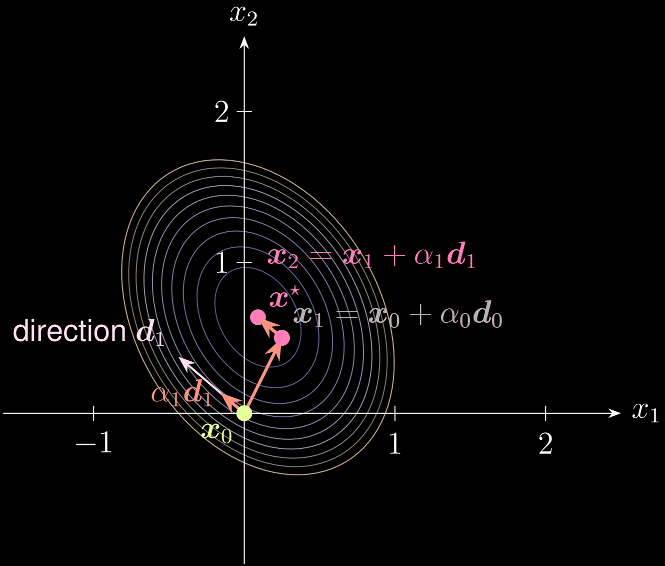
Conjugate Gradient Method



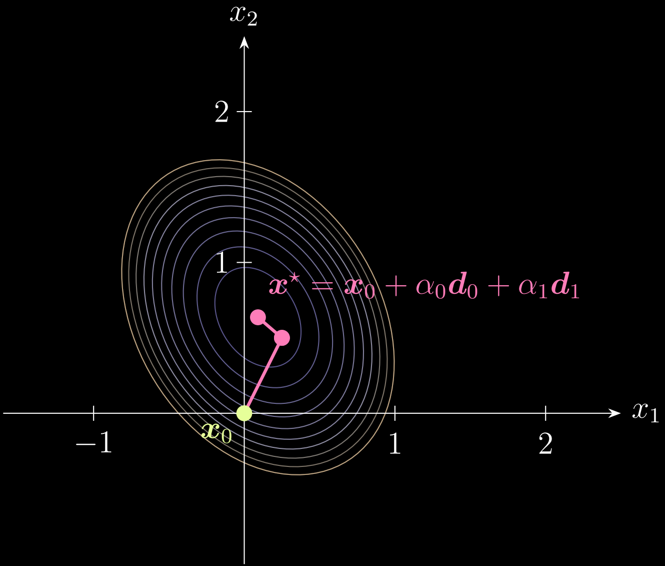
$$\boldsymbol{d}_1 = \begin{bmatrix} -7/16 \\ 3/8 \end{bmatrix}$$

$$\alpha_1 = 4/11$$

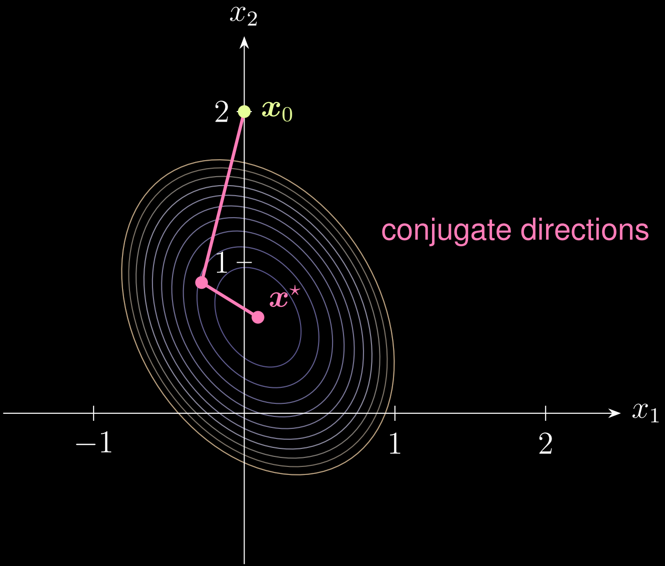
Conjugate Gradient Method



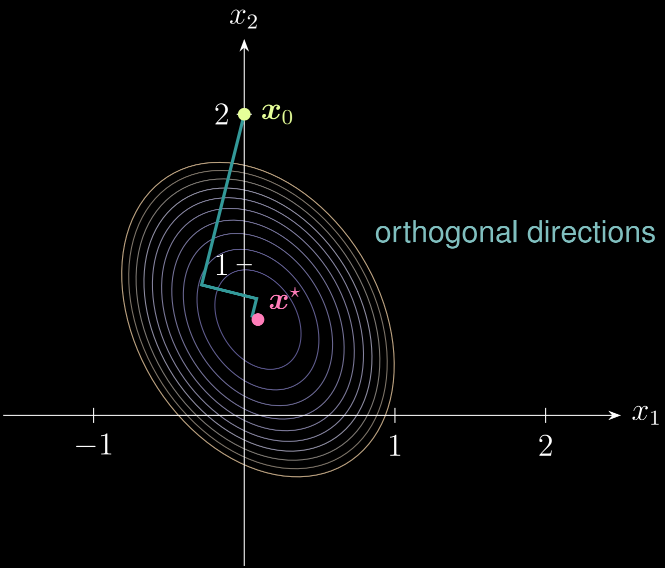
Conjugate Gradient Method



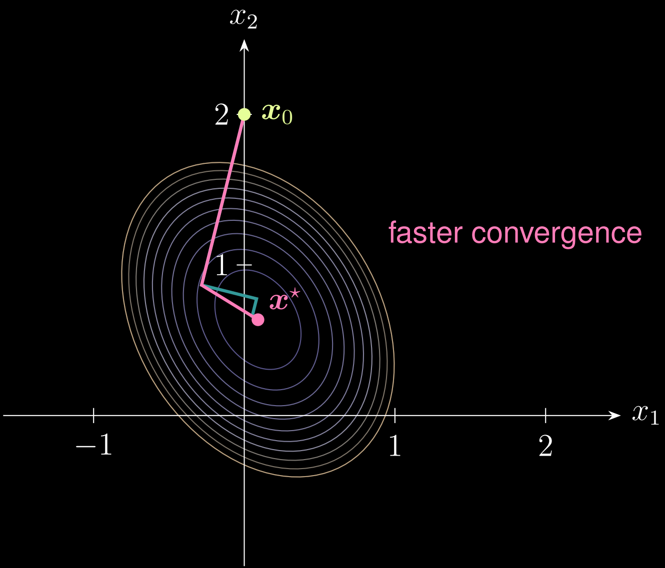
Conjugate Gradient Method



Steepest Gradient Descent Method

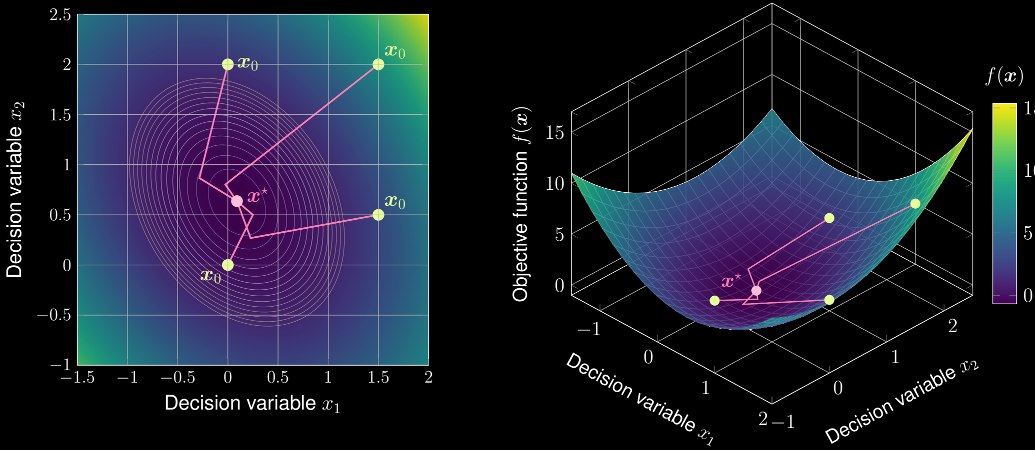


Conjugate Directions vs. Orthogonal Directions



Conjugate Gradient Method for Solving Quadratic Optimization

$$\min_{\boldsymbol{x}\in\mathbb{R}^2} \quad \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}^\top \boldsymbol{x} \quad \text{with} \quad \boldsymbol{A} = \begin{bmatrix} 4 & 1 \\ 1 & 3 \end{bmatrix} \quad \text{and} \quad \boldsymbol{b} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$



Thanks for your attention!

About me:

- Homepage: https://xinychen.github.io
- How to reach me: chenxy346@gmail.com