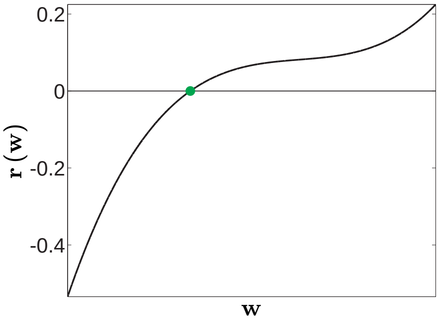
Illustration of the Newton method

Note the different notation used in these slides compared to the Lecture notes:

	Lecture notes	**These slides**
Optimization variables	x	w
Objective function	$f(x)$	$\Phi(w)$
Inequality constraint function	$g(x)$	$h(w)$
Linear constraint function	$h(x)$	$g(w)$

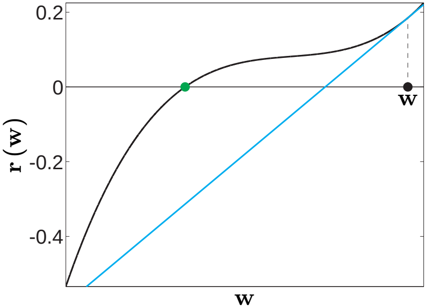
Core idea

Goal: solve $\mathbf{r}(\mathbf{w}) = 0$... **how ?!?**



Core idea

Goal: solve $\mathbf{r}(\mathbf{w}) = 0$... **how ?!?**

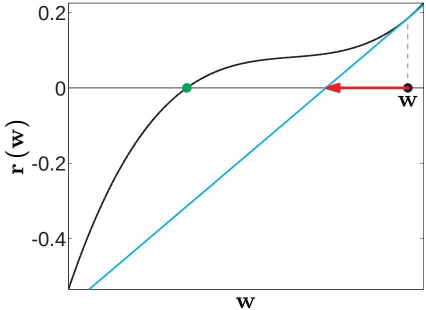


Key idea: guess $\mathbf{w}$, iterate the linear model:

$$\mathbf{r}(\mathbf{w} + \Delta\mathbf{w}) \approx \mathbf{r}(\mathbf{w}) + \nabla\mathbf{r}(\mathbf{w})^\top \Delta\mathbf{w} = 0$$

Core idea

Goal: solve $\mathbf{r}\left(\mathbf{w}\right) = 0$... **how ?!?**

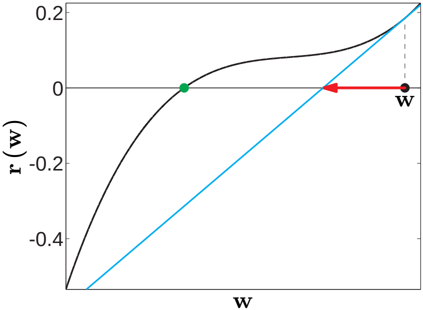


Key idea: guess $\mathbf{w}$, iterate the linear model:

$$\mathbf{r}\left(\mathbf{w} + \Delta\mathbf{w}\right) \approx \mathbf{r}\left(\mathbf{w}\right) + \nabla\mathbf{r}\left(\mathbf{w}\right)^\top \Delta\mathbf{w} = 0$$

Core idea

Goal: solve $\mathbf{r}(\mathbf{w}) = 0$... **how ?!?**



Key idea: guess $\mathbf{w}$, iterate the linear model:

$$\mathbf{r}(\mathbf{w} + \Delta\mathbf{w}) \approx \mathbf{r}(\mathbf{w}) + \nabla\mathbf{r}(\mathbf{w})^\top \Delta\mathbf{w} = 0$$

Algorithm: Newton method

Input: $\mathbf{w}$, Tol

while $\|\mathbf{r}(\mathbf{w})\| \geq \text{tol}$ **do**

Compute

$$\mathbf{r}(\mathbf{w}) \quad \text{and} \quad \nabla\mathbf{r}(\mathbf{w})$$

Compute the Newton direction

$$\nabla\mathbf{r}(\mathbf{w})^\top \Delta\mathbf{w} = -\mathbf{r}(\mathbf{w})$$

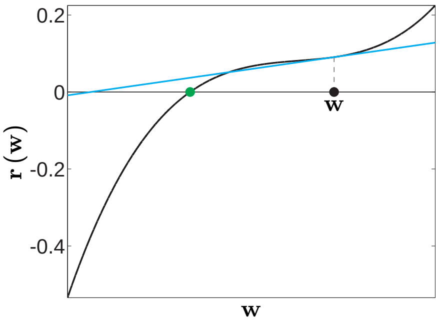
Newton step

$$\mathbf{w} \leftarrow \mathbf{w} + \Delta\mathbf{w}$$

return $\mathbf{w}$

Core idea

Goal: solve $\mathbf{r}(\mathbf{w}) = 0$... **how ?!?**



Key idea: guess $\mathbf{w}$, iterate the linear model:

$$\mathbf{r}(\mathbf{w} + \Delta\mathbf{w}) \approx \mathbf{r}(\mathbf{w}) + \nabla\mathbf{r}(\mathbf{w})^\top \Delta\mathbf{w} = 0$$

Algorithm: Newton method

Input: $\mathbf{w}$, Tol

while $\|\mathbf{r}(\mathbf{w})\| \geq \text{tol}$ **do**

Compute

$$\mathbf{r}(\mathbf{w}) \quad \text{and} \quad \nabla\mathbf{r}(\mathbf{w})$$

Compute the Newton direction

$$\nabla\mathbf{r}(\mathbf{w})^\top \Delta\mathbf{w} = -\mathbf{r}(\mathbf{w})$$

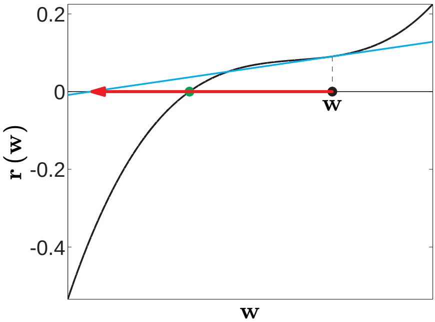
Newton step

$$\mathbf{w} \leftarrow \mathbf{w} + \Delta\mathbf{w}$$

return $\mathbf{w}$

Core idea

Goal: solve $\mathbf{r}(\mathbf{w}) = 0$... **how ?!?**



Key idea: guess $\mathbf{w}$, iterate the linear model:

$$\mathbf{r}(\mathbf{w} + \Delta\mathbf{w}) \approx \mathbf{r}(\mathbf{w}) + \nabla\mathbf{r}(\mathbf{w})^\top \Delta\mathbf{w} = 0$$

Algorithm: Newton method

Input: $\mathbf{w}$, Tol

while $\|\mathbf{r}(\mathbf{w})\| \geq \text{tol}$ **do**

Compute

$$\mathbf{r}(\mathbf{w}) \quad \text{and} \quad \nabla\mathbf{r}(\mathbf{w})$$

Compute the Newton direction

$$\nabla\mathbf{r}(\mathbf{w})^\top \Delta\mathbf{w} = -\mathbf{r}(\mathbf{w})$$

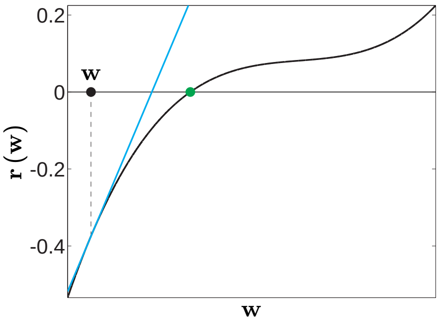
Newton step

$$\mathbf{w} \leftarrow \mathbf{w} + \Delta\mathbf{w}$$

return $\mathbf{w}$

Core idea

Goal: solve $\mathbf{r}(\mathbf{w}) = 0$... **how ?!?**



Key idea: guess $\mathbf{w}$, iterate the linear model:

$$\mathbf{r}(\mathbf{w} + \Delta\mathbf{w}) \approx \mathbf{r}(\mathbf{w}) + \nabla\mathbf{r}(\mathbf{w})^\top \Delta\mathbf{w} = 0$$

Algorithm: Newton method

Input: $\mathbf{w}$, Tol

while $\|\mathbf{r}(\mathbf{w})\| \geq \text{tol}$ **do**

Compute

$$\mathbf{r}(\mathbf{w}) \quad \text{and} \quad \nabla\mathbf{r}(\mathbf{w})$$

Compute the Newton direction

$$\nabla\mathbf{r}(\mathbf{w})^\top \Delta\mathbf{w} = -\mathbf{r}(\mathbf{w})$$

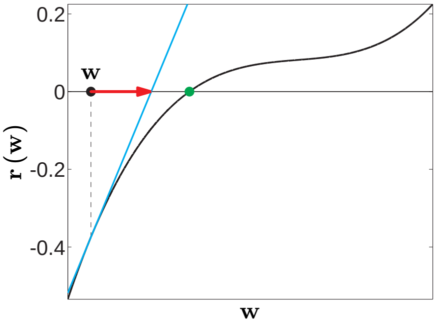
Newton step

$$\mathbf{w} \leftarrow \mathbf{w} + \Delta\mathbf{w}$$

return $\mathbf{w}$

Core idea

Goal: solve $\mathbf{r}(\mathbf{w}) = 0$... **how ?!?**



Key idea: guess $\mathbf{w}$, iterate the linear model:

$$\mathbf{r}(\mathbf{w} + \Delta\mathbf{w}) \approx \mathbf{r}(\mathbf{w}) + \nabla\mathbf{r}(\mathbf{w})^\top \Delta\mathbf{w} = 0$$

Algorithm: Newton method

Input: $\mathbf{w}$, Tol

while $\|\mathbf{r}(\mathbf{w})\| \geq \mathrm{tol}$ **do**

Compute

$$\mathbf{r}(\mathbf{w}) \quad \text{and} \quad \nabla\mathbf{r}(\mathbf{w})$$

Compute the Newton direction

$$\nabla\mathbf{r}(\mathbf{w})^\top \Delta\mathbf{w} = -\mathbf{r}(\mathbf{w})$$

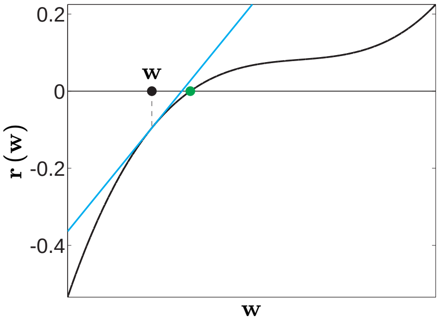
Newton step

$$\mathbf{w} \leftarrow \mathbf{w} + \Delta\mathbf{w}$$

return $\mathbf{w}$

Core idea

Goal: solve $\mathbf{r}(\mathbf{w}) = 0$... **how ?!?**



Key idea: guess $\mathbf{w}$, iterate the linear model:

$$\mathbf{r}(\mathbf{w} + \Delta\mathbf{w}) \approx \mathbf{r}(\mathbf{w}) + \nabla\mathbf{r}(\mathbf{w})^\top \Delta\mathbf{w} = 0$$

Algorithm: Newton method

Input: $\mathbf{w}$, Tol

while $\|\mathbf{r}(\mathbf{w})\| \geq \text{tol}$ **do**

Compute

$$\mathbf{r}(\mathbf{w}) \quad \text{and} \quad \nabla\mathbf{r}(\mathbf{w})$$

Compute the Newton direction

$$\nabla\mathbf{r}(\mathbf{w})^\top \Delta\mathbf{w} = -\mathbf{r}(\mathbf{w})$$

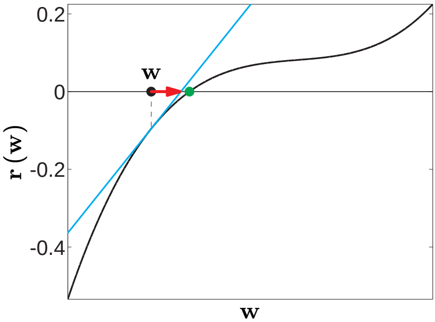
Newton step

$$\mathbf{w} \leftarrow \mathbf{w} + \Delta\mathbf{w}$$

return $\mathbf{w}$

Core idea

Goal: solve $\mathbf{r}(\mathbf{w}) = 0$... **how ?!?**



Key idea: guess $\mathbf{w}$, iterate the linear model:

$$\mathbf{r}(\mathbf{w} + \Delta\mathbf{w}) \approx \mathbf{r}(\mathbf{w}) + \nabla\mathbf{r}(\mathbf{w})^\top \Delta\mathbf{w} = 0$$

Algorithm: Newton method

Input: $\mathbf{w}$, Tol

while $\|\mathbf{r}(\mathbf{w})\| \geq \text{tol}$ **do**

Compute

$$\mathbf{r}(\mathbf{w}) \quad \text{and} \quad \nabla\mathbf{r}(\mathbf{w})$$

Compute the Newton direction

$$\nabla\mathbf{r}(\mathbf{w})^\top \Delta\mathbf{w} = -\mathbf{r}(\mathbf{w})$$

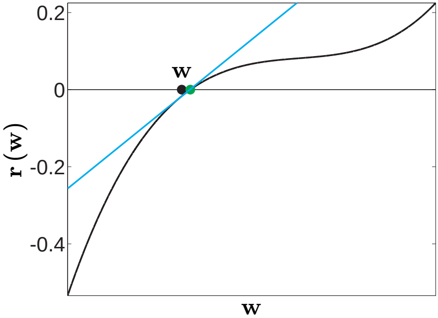
Newton step

$$\mathbf{w} \leftarrow \mathbf{w} + \Delta\mathbf{w}$$

return $\mathbf{w}$

Core idea

Goal: solve $\mathbf{r}(\mathbf{w}) = 0$... **how ?!?**



Key idea: guess $\mathbf{w}$, iterate the linear model:

$$\mathbf{r}(\mathbf{w} + \Delta\mathbf{w}) \approx \mathbf{r}(\mathbf{w}) + \nabla\mathbf{r}(\mathbf{w})^\top \Delta\mathbf{w} = 0$$

Algorithm: Newton method

Input: $\mathbf{w}$, Tol

while $\|\mathbf{r}(\mathbf{w})\| \geq \mathrm{tol}$ **do**

Compute

$$\mathbf{r}(\mathbf{w}) \quad \text{and} \quad \nabla\mathbf{r}(\mathbf{w})$$

Compute the Newton direction

$$\nabla\mathbf{r}(\mathbf{w})^\top \Delta\mathbf{w} = -\mathbf{r}(\mathbf{w})$$

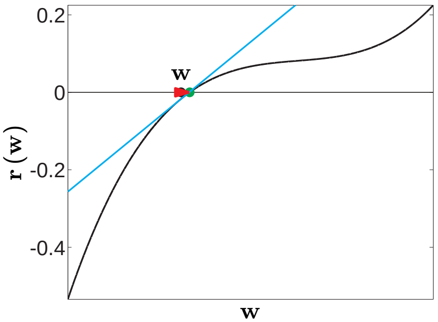
Newton step

$$\mathbf{w} \leftarrow \mathbf{w} + \Delta\mathbf{w}$$

return $\mathbf{w}$

Core idea

Goal: solve $\mathbf{r}(\mathbf{w}) = 0$... **how ?!?**



Key idea: guess $\mathbf{w}$, iterate the linear model:

$$\mathbf{r}(\mathbf{w} + \Delta\mathbf{w}) \approx \mathbf{r}(\mathbf{w}) + \nabla\mathbf{r}(\mathbf{w})^\top \Delta\mathbf{w} = 0$$

Algorithm: Newton method

Input: $\mathbf{w}$, Tol

while $\|\mathbf{r}(\mathbf{w})\| \geq \mathrm{tol}$ **do**

Compute

$$\mathbf{r}(\mathbf{w}) \quad \text{and} \quad \nabla\mathbf{r}(\mathbf{w})$$

Compute the Newton direction

$$\nabla\mathbf{r}(\mathbf{w})^\top \Delta\mathbf{w} = -\mathbf{r}(\mathbf{w})$$

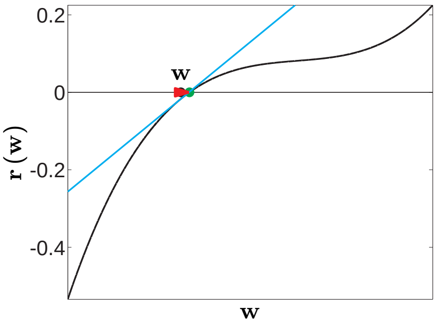
Newton step

$$\mathbf{w} \leftarrow \mathbf{w} + \Delta\mathbf{w}$$

return $\mathbf{w}$

Core idea

Goal: solve $\mathbf{r}(\mathbf{w}) = 0$... **how ?!?**



Key idea: guess $\mathbf{w}$, iterate the linear model:

$$\mathbf{r}(\mathbf{w} + \Delta\mathbf{w}) \approx \mathbf{r}(\mathbf{w}) + \nabla\mathbf{r}(\mathbf{w})^\top \Delta\mathbf{w} = 0$$

Algorithm: Newton method

Input: $\mathbf{w}$, Tol

while $\|\mathbf{r}(\mathbf{w})\| \geq \text{tol}$ **do**

  Compute

$$\mathbf{r}(\mathbf{w}) \quad \text{and} \quad \nabla\mathbf{r}(\mathbf{w})$$

  Compute the Newton direction

$$\nabla\mathbf{r}(\mathbf{w})^\top \Delta\mathbf{w} = -\mathbf{r}(\mathbf{w})$$

  Newton step

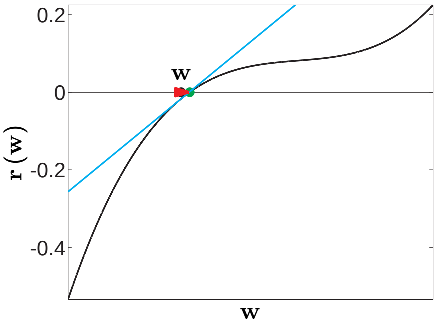
$$\mathbf{w} \leftarrow \mathbf{w} + \Delta\mathbf{w}$$

return $\mathbf{w}$

- This is a **full-step** Newton iteration

Core idea

Goal: solve $\mathbf{r}(\mathbf{w}) = 0$... **how ?!?**



Key idea: guess $\mathbf{w}$, iterate the linear model:

$$\mathbf{r}(\mathbf{w} + \Delta\mathbf{w}) \approx \mathbf{r}(\mathbf{w}) + \nabla\mathbf{r}(\mathbf{w})^\top \Delta\mathbf{w} = 0$$

Algorithm: Newton method

Input: $\mathbf{w}$, Tol

while $\|\mathbf{r}(\mathbf{w})\| \geq \text{tol}$ **do**

Compute

$$\mathbf{r}(\mathbf{w}) \quad \text{and} \quad \nabla\mathbf{r}(\mathbf{w})$$

Compute the Newton direction

$$\nabla\mathbf{r}(\mathbf{w})^\top \Delta\mathbf{w} = -\mathbf{r}(\mathbf{w})$$

Newton step, $t \in]0, 1]$

$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}$$

return $\mathbf{w}$

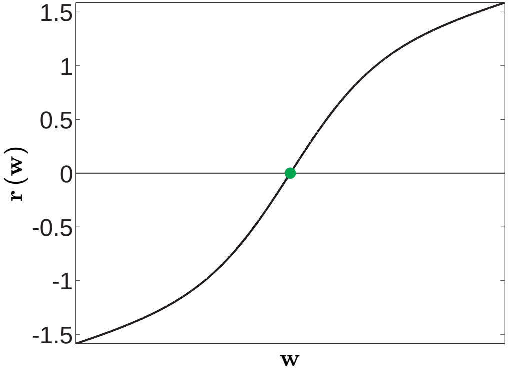
- This is a **full-step** Newton iteration
- Reduced steps are often needed

Why reduced steps ?

Newton step with $t \in]0, 1]$:

$$\nabla \mathbf{r}(\mathbf{w}) \, \Delta \mathbf{w} = -\mathbf{r}(\mathbf{w})$$

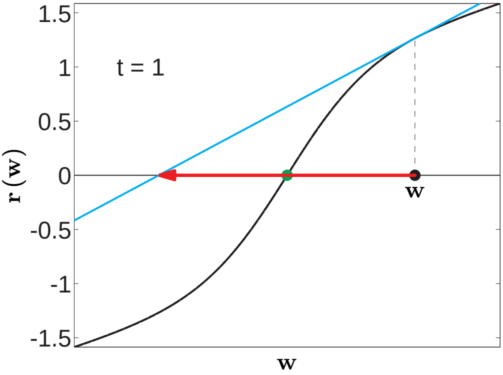
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta \mathbf{w}$$



Why reduced steps ?

Newton step with $t \in]0, 1]$:

$$\nabla \mathbf{r}(\mathbf{w})\, \Delta\mathbf{w} = -\mathbf{r}(\mathbf{w})$$
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}$$

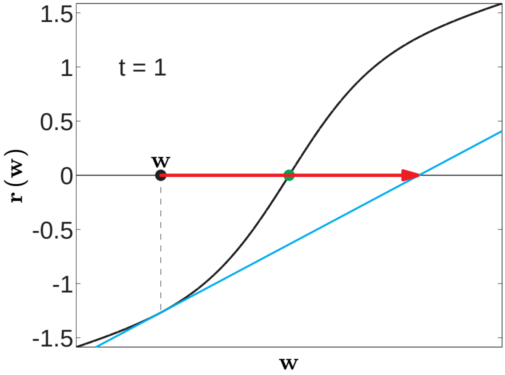


Why reduced steps ?

Newton step with $t \in]0, 1]$:

$$\nabla \mathbf{r}(\mathbf{w})\,\Delta\mathbf{w} = -\mathbf{r}(\mathbf{w})$$

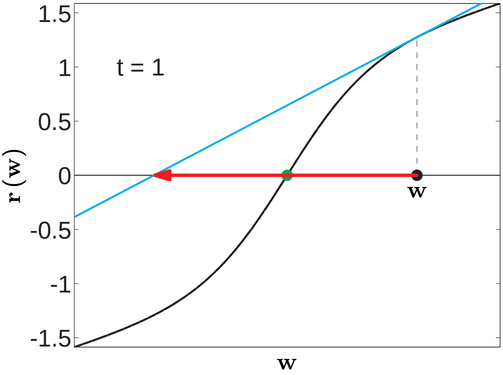
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}$$



Why reduced steps ?

Newton step with $t \in]0, 1]$:

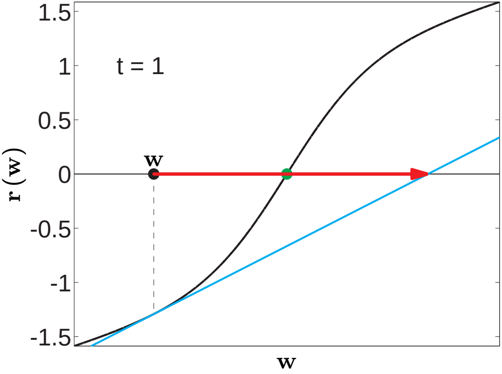
$$\nabla \mathbf{r}(\mathbf{w})\, \Delta\mathbf{w} = -\mathbf{r}(\mathbf{w})$$
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}$$



Why reduced steps ?

Newton step with $t \in]0, 1]$:

$$\nabla \mathbf{r}(\mathbf{w})\, \Delta\mathbf{w} = -\mathbf{r}(\mathbf{w})$$
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}$$

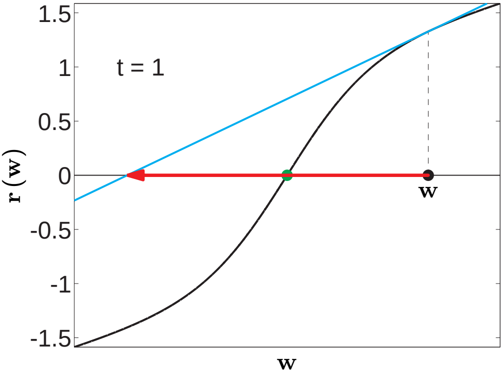


Why reduced steps ?

Newton step with $t \in]0, 1]$:

$$\nabla \mathbf{r}(\mathbf{w}) \Delta \mathbf{w} = -\mathbf{r}(\mathbf{w})$$

$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta \mathbf{w}$$

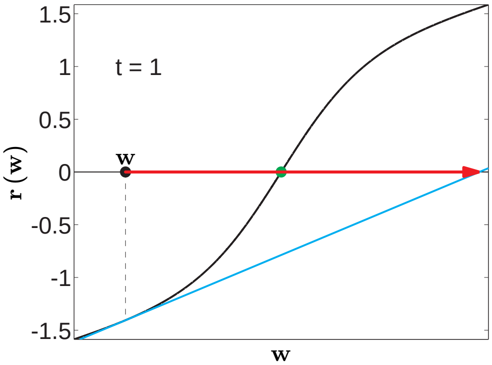


Why reduced steps ?

Newton step with $t \in]0, 1]$:

$$\nabla \mathbf{r}(\mathbf{w})\, \Delta\mathbf{w} = -\mathbf{r}(\mathbf{w})$$

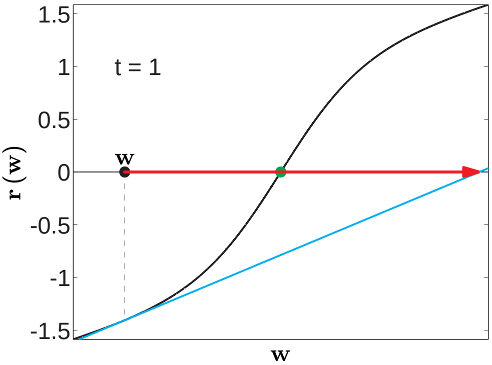
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}$$



Why reduced steps ?

Newton step with $t \in]0,\, 1]$:

$$\nabla \mathbf{r}(\mathbf{w})\, \Delta\mathbf{w} = -\mathbf{r}(\mathbf{w})$$
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}$$

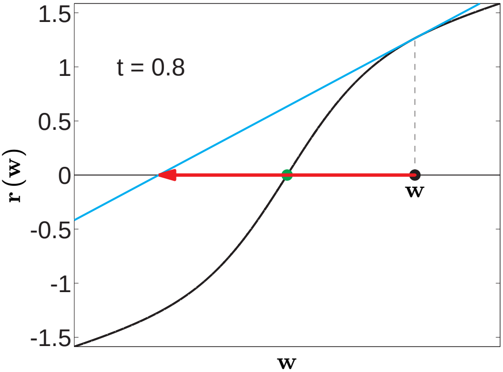


The full-step Newton iteration can be unstable !!

Why reduced steps ?

Newton step with $t \in \,]0, 1]$:

$$\nabla \mathbf{r}(\mathbf{w})\, \Delta \mathbf{w} = -\mathbf{r}(\mathbf{w})$$
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta \mathbf{w}$$

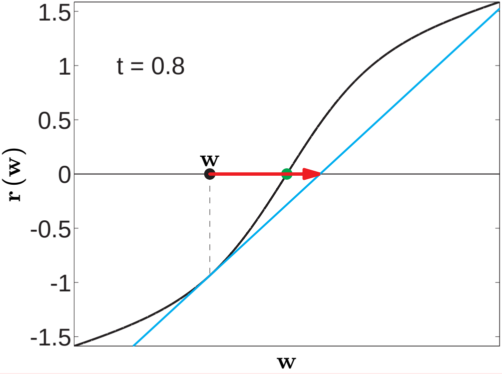


The full-step Newton iteration can be unstable !!

Why reduced steps ?

Newton step with $t \in]0, 1]$:

$$\nabla \mathbf{r}(\mathbf{w}) \Delta\mathbf{w} = -\mathbf{r}(\mathbf{w})$$
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}$$

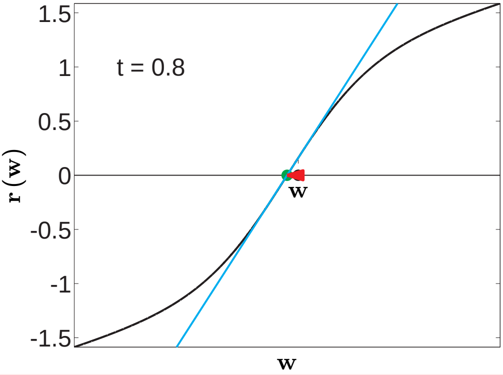


The full-step Newton iteration can be unstable !!

Why reduced steps ?

Newton step with $t \in]0, 1]$:

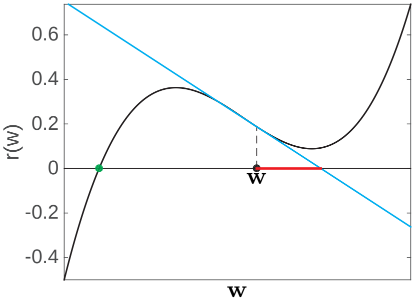
$$\nabla \mathbf{r}(\mathbf{w}) \, \Delta \mathbf{w} = -\mathbf{r}(\mathbf{w})$$
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta \mathbf{w}$$



The full-step Newton iteration can be unstable !!

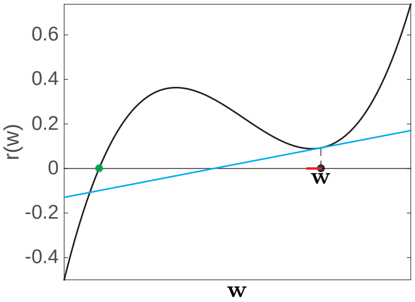
But still, Newton can fail...

Solve $\mathbf{r}\left(\mathbf{w}\right) = 0$



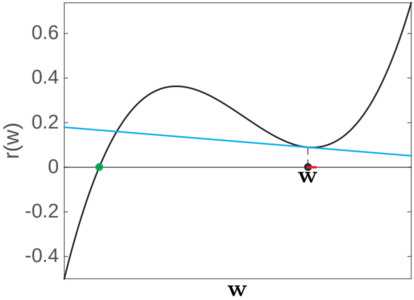
But still, Newton can fail...

Solve $\mathbf{r}(\mathbf{w}) = 0$



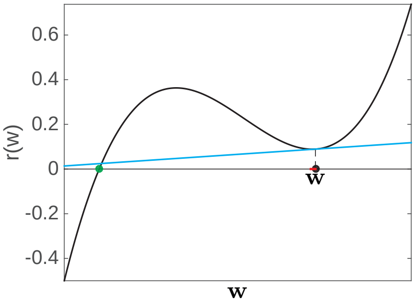
But still, Newton can fail...

Solve $\mathbf{r}(\mathbf{w}) = 0$



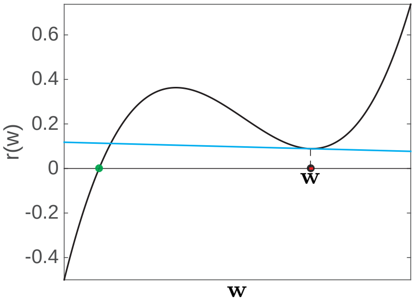
But still, Newton can fail...

Solve $\mathbf{r}(\mathbf{w}) = 0$



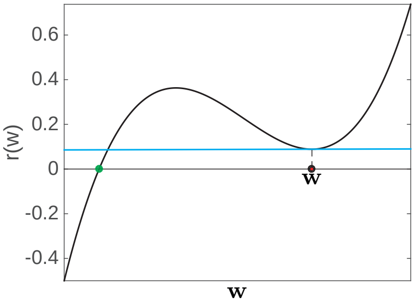
But still, Newton can fail...

Solve $\mathbf{r}(\mathbf{w}) = 0$



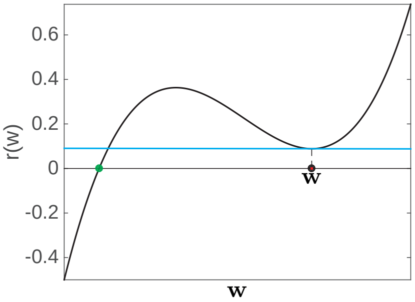
But still, Newton can fail...

Solve $\mathbf{r}(\mathbf{w}) = 0$



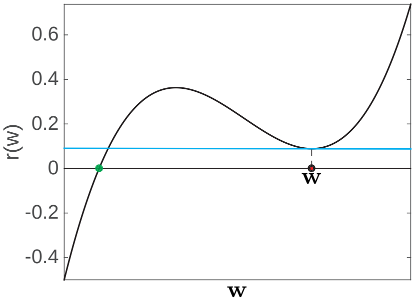
But still, Newton can fail...

Solve $\mathbf{r}(\mathbf{w}) = 0$



But still, Newton can fail...

Solve $\mathbf{r}(\mathbf{w}) = 0$



Newton stops with

$$\mathbf{r}(\mathbf{w}) \neq 0 \text{ and } \nabla \mathbf{r}(\mathbf{w}) \text{ singular}$$

i.e. the Newton direction $\Delta \mathbf{w}$ given by

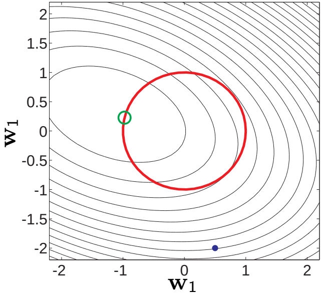
$$\nabla \mathbf{r}(\mathbf{w}) \Delta \mathbf{w} = -\mathbf{r}(\mathbf{w})$$

is undefined

Newton Iteration for Optimization - Example

$$\min_{\mathbf{w}} \; \frac{1}{2}\mathbf{w}^\top \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \mathbf{w} + \mathbf{w}^\top \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t. } g(\mathbf{w}) = \mathbf{w}^\top \mathbf{w} - 1 = 0$$



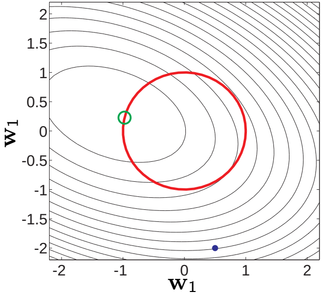
Newton Iteration for Optimization - Example

Iterate:

$$\begin{bmatrix} H & \nabla \mathbf{g} \\ \nabla \mathbf{g}^\top & 0 \end{bmatrix} \begin{bmatrix} \Delta \mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = - \begin{bmatrix} \nabla \Phi \\ \mathbf{g} \end{bmatrix}$$

$$\min_{\mathbf{w}} \ \frac{1}{2}\mathbf{w}^\top \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \mathbf{w} + \mathbf{w}^\top \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t. } g(\mathbf{w}) = \mathbf{w}^\top \mathbf{w} - 1 = 0$$



Newton Iteration for Optimization - Example

Iterate:

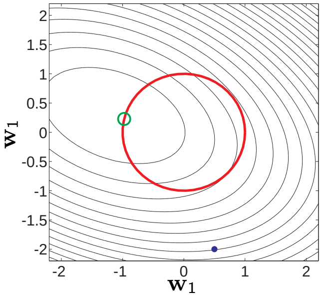
$$\begin{bmatrix} H & \nabla \mathbf{g} \\ \nabla \mathbf{g}^\top & 0 \end{bmatrix} \begin{bmatrix} \Delta \mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = - \begin{bmatrix} \nabla \Phi \\ \mathbf{g} \end{bmatrix}$$

with:

$$\nabla g(\mathbf{w}) = 2\mathbf{w} = \begin{bmatrix} 2\mathbf{w}_1 \\ 2\mathbf{w}_2 \end{bmatrix}$$

$$\min_{\mathbf{w}} \ \frac{1}{2}\mathbf{w}^\top \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \mathbf{w} + \mathbf{w}^\top \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t. } g(\mathbf{w}) = \mathbf{w}^\top \mathbf{w} - 1 = 0$$



Newton Iteration for Optimization - Example

Iterate:

$$\begin{bmatrix} H & \nabla \mathbf{g} \\ \nabla \mathbf{g}^\top & 0 \end{bmatrix} \begin{bmatrix} \Delta \mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = - \begin{bmatrix} \nabla \Phi \\ \mathbf{g} \end{bmatrix}$$

with:

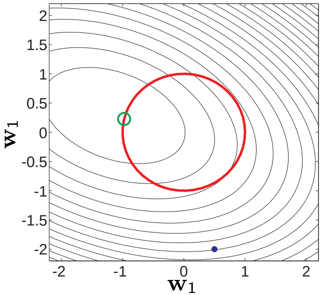
$$\nabla g(\mathbf{w}) = 2\mathbf{w} = \begin{bmatrix} 2\mathbf{w}_1 \\ 2\mathbf{w}_2 \end{bmatrix}$$

$$\mathcal{L}(\mathbf{w}, \lambda) = \Phi(\mathbf{w}) + \lambda g(\mathbf{w})$$

$$\nabla_{\mathbf{w}} \mathcal{L}(\mathbf{w}, \lambda) = \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \mathbf{w} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 2\lambda \mathbf{w}$$

$$\min_{\mathbf{w}} \ \frac{1}{2} \mathbf{w}^\top \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \mathbf{w} + \mathbf{w}^\top \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t. } g(\mathbf{w}) = \mathbf{w}^\top \mathbf{w} - 1 = 0$$



Newton Iteration for Optimization - Example

Iterate:

$$\begin{bmatrix} H & \nabla \mathbf{g} \\ \nabla \mathbf{g}^\top & 0 \end{bmatrix} \begin{bmatrix} \Delta \mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = - \begin{bmatrix} \nabla \Phi \\ \mathbf{g} \end{bmatrix}$$

with:

$$\nabla g(\mathbf{w}) = 2\mathbf{w} = \begin{bmatrix} 2\mathbf{w}_1 \\ 2\mathbf{w}_2 \end{bmatrix}$$

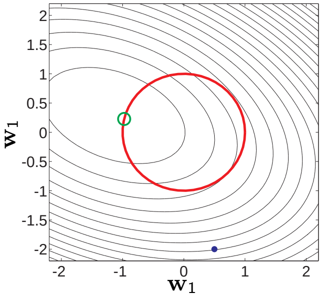
$$\mathcal{L}(\mathbf{w}, \lambda) = \Phi(\mathbf{w}) + \lambda g(\mathbf{w})$$

$$\nabla_{\mathbf{w}} \mathcal{L}(\mathbf{w}, \lambda) = \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \mathbf{w} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 2\lambda \mathbf{w}$$

$$H(\mathbf{w}, \lambda) = \begin{bmatrix} 2 + 2\lambda & 1 \\ 1 & 4 + 2\lambda \end{bmatrix}$$

$$\min_{\mathbf{w}} \ \frac{1}{2} \mathbf{w}^\top \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \mathbf{w} + \mathbf{w}^\top \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t. } g(\mathbf{w}) = \mathbf{w}^\top \mathbf{w} - 1 = 0$$



Newton Iteration for Optimization - Example

Iterate:

$$\begin{bmatrix} H & \nabla \mathbf{g} \\ \nabla \mathbf{g}^\top & 0 \end{bmatrix} \begin{bmatrix} \Delta \mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = - \begin{bmatrix} \nabla \Phi \\ \mathbf{g} \end{bmatrix}$$

with:

$$\nabla g(\mathbf{w}) = 2\mathbf{w} = \begin{bmatrix} 2\mathbf{w}_1 \\ 2\mathbf{w}_2 \end{bmatrix}$$

$$\mathcal{L}(\mathbf{w}, \lambda) = \Phi(\mathbf{w}) + \lambda g(\mathbf{w})$$

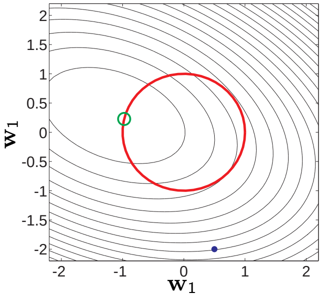
$$\nabla_{\mathbf{w}} \mathcal{L}(\mathbf{w}, \lambda) = \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \mathbf{w} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 2\lambda \mathbf{w}$$

$$H(\mathbf{w}, \lambda) = \begin{bmatrix} 2 + 2\lambda & 1 \\ 1 & 4 + 2\lambda \end{bmatrix}$$

$$\nabla \Phi(\mathbf{w}) = \begin{bmatrix} 2\mathbf{w}_1 + \mathbf{w}_2 + 1 \\ \mathbf{w}_1 + 4\mathbf{w}_2 \end{bmatrix}$$

$$\min_{\mathbf{w}} \frac{1}{2} \mathbf{w}^\top \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \mathbf{w} + \mathbf{w}^\top \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t. } g(\mathbf{w}) = \mathbf{w}^\top \mathbf{w} - 1 = 0$$



Newton Iteration for Optimization - Example

Algorithm: Newton method

Input: guess $\mathbf{w}$, $\boldsymbol{\lambda}$
while $\|\nabla\mathcal{L}\|$ *or* $\|\mathbf{g}\| \geq \text{tol}$ **do**

Compute

$$H(\mathbf{w}, \boldsymbol{\lambda}), \ \nabla\mathbf{g}(\mathbf{w}), \ \nabla\Phi(\mathbf{w}), \ \mathbf{g}(\mathbf{w})$$

Compute Newton direction

$$\begin{bmatrix} H & \nabla\mathbf{g} \\ \nabla\mathbf{g}^\top & 0 \end{bmatrix} \begin{bmatrix} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = -\begin{bmatrix} \nabla\Phi \\ \mathbf{g} \end{bmatrix}$$

$$\Delta\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ - \boldsymbol{\lambda}$$

Compute Newton step, $t \in]0, 1]$

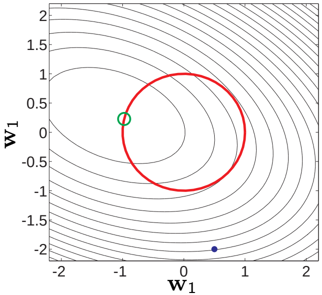
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + t\Delta\boldsymbol{\lambda}$$

return $\mathbf{w}$, $\boldsymbol{\lambda}$

$$\min_{\mathbf{w}} \ \frac{1}{2}\mathbf{w}^\top \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \mathbf{w} + \mathbf{w}^\top \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t.} \ g(\mathbf{w}) = \mathbf{w}^\top\mathbf{w} - 1 = 0$$

Guess $\lambda = 0$, step $t = 1$



Newton Iteration for Optimization - Example

Algorithm: Newton method

Input: guess $\mathbf{w}$, $\boldsymbol{\lambda}$

while $\|\nabla\mathcal{L}\|$ *or* $\|\mathbf{g}\| \geq \text{tol}$ **do**

Compute

$$H(\mathbf{w}, \boldsymbol{\lambda}), \nabla \mathbf{g}(\mathbf{w}), \nabla \Phi(\mathbf{w}), \mathbf{g}(\mathbf{w})$$

Compute Newton direction

$$\begin{bmatrix} H & \nabla \mathbf{g} \\ \nabla \mathbf{g}^\top & 0 \end{bmatrix} \begin{bmatrix} \Delta \mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = - \begin{bmatrix} \nabla \Phi \\ \mathbf{g} \end{bmatrix}$$

$$\Delta \boldsymbol{\lambda} = \boldsymbol{\lambda}^+ - \boldsymbol{\lambda}$$

Compute Newton step, $t \in]0, 1]$

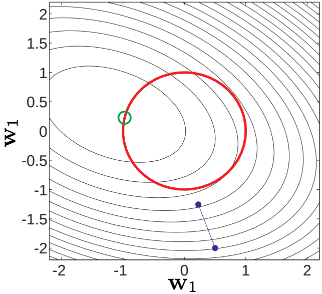
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + t\Delta\boldsymbol{\lambda}$$

return $\mathbf{w}$, $\boldsymbol{\lambda}$

$$\min_{\mathbf{w}} \ \frac{1}{2}\mathbf{w}^\top \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \mathbf{w} + \mathbf{w}^\top \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t. } g(\mathbf{w}) = \mathbf{w}^\top \mathbf{w} - 1 = 0$$

Guess $\lambda = 0$, step $t = 1$



Newton Iteration for Optimization - Example

Algorithm: Newton method

Input: guess $\mathbf{w}$, $\boldsymbol{\lambda}$

while $\|\nabla\mathcal{L}\|$ *or* $\|\mathbf{g}\| \geq \mathrm{tol}$ **do**

Compute

$$H(\mathbf{w},\boldsymbol{\lambda}),\ \nabla\mathbf{g}(\mathbf{w}),\ \nabla\Phi(\mathbf{w}),\ \mathbf{g}(\mathbf{w})$$

Compute Newton direction

$$\begin{bmatrix} H & \nabla\mathbf{g} \\ \nabla\mathbf{g}^\top & 0 \end{bmatrix}\begin{bmatrix} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = -\begin{bmatrix} \nabla\Phi \\ \mathbf{g} \end{bmatrix}$$

$$\Delta\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ - \boldsymbol{\lambda}$$

Compute Newton step, $t \in]0,1]$

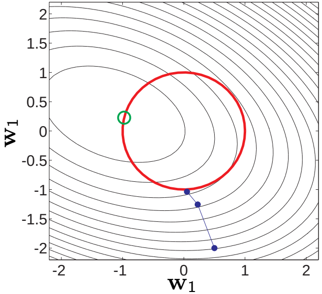
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + t\Delta\boldsymbol{\lambda}$$

return $\mathbf{w}$, $\boldsymbol{\lambda}$

$$\min_{\mathbf{w}} \ \frac{1}{2}\mathbf{w}^\top \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}\mathbf{w} + \mathbf{w}^\top \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t.} \ g(\mathbf{w}) = \mathbf{w}^\top\mathbf{w} - 1 = 0$$

Guess $\lambda = 0$, step $t = 1$



Newton Iteration for Optimization - Example

Algorithm: Newton method

Input: guess $\mathbf{w}$, $\boldsymbol{\lambda}$

while $\|\nabla\mathcal{L}\|$ *or* $\|\mathbf{g}\| \geq \mathrm{tol}$ **do**

Compute

$$H(\mathbf{w},\boldsymbol{\lambda}),\ \nabla\mathbf{g}(\mathbf{w}),\ \nabla\Phi(\mathbf{w}),\ \mathbf{g}(\mathbf{w})$$

Compute Newton direction

$$\begin{bmatrix} H & \nabla\mathbf{g} \\ \nabla\mathbf{g}^\top & 0 \end{bmatrix}\begin{bmatrix} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = -\begin{bmatrix} \nabla\Phi \\ \mathbf{g} \end{bmatrix}$$

$$\Delta\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ - \boldsymbol{\lambda}$$

Compute Newton step, $t \in]0,1]$

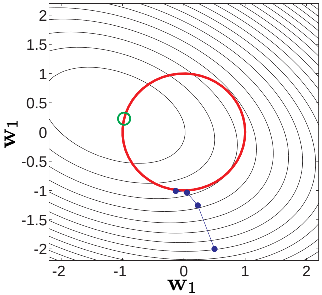
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + t\Delta\boldsymbol{\lambda}$$

return $\mathbf{w}$, $\boldsymbol{\lambda}$

$$\min_{\mathbf{w}} \ \frac{1}{2}\mathbf{w}^\top\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}\mathbf{w} + \mathbf{w}^\top\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t.} \ g(\mathbf{w}) = \mathbf{w}^\top\mathbf{w} - 1 = 0$$

Guess $\lambda = 0$, step $t = 1$



Newton Iteration for Optimization - Example

Algorithm: Newton method

Input: guess $\mathbf{w}$, $\boldsymbol{\lambda}$

while $\|\nabla\mathcal{L}\|$ *or* $\|\mathbf{g}\| \geq \mathrm{tol}$ **do**

Compute

$$H(\mathbf{w},\boldsymbol{\lambda}),\ \nabla\mathbf{g}(\mathbf{w}),\ \nabla\Phi(\mathbf{w}),\ \mathbf{g}(\mathbf{w})$$

Compute Newton direction

$$\begin{bmatrix} H & \nabla\mathbf{g} \\ \nabla\mathbf{g}^\top & 0 \end{bmatrix} \begin{bmatrix} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = - \begin{bmatrix} \nabla\Phi \\ \mathbf{g} \end{bmatrix}$$

$$\Delta\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ - \boldsymbol{\lambda}$$

Compute Newton step, $t \in]0,1]$

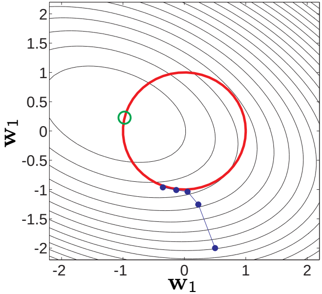
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + t\Delta\boldsymbol{\lambda}$$

return $\mathbf{w}$, $\boldsymbol{\lambda}$

$$\min_{\mathbf{w}} \ \frac{1}{2}\mathbf{w}^\top \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \mathbf{w} + \mathbf{w}^\top \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t. } g(\mathbf{w}) = \mathbf{w}^\top\mathbf{w} - 1 = 0$$

Guess $\lambda = 0$, step $t = 1$



Newton Iteration for Optimization - Example

Algorithm: Newton method

Input: guess $\mathbf{w}$, $\boldsymbol{\lambda}$

while $\|\nabla\mathcal{L}\|$ *or* $\|\mathbf{g}\| \geq \text{tol}$ **do**

Compute

$$H(\mathbf{w}, \boldsymbol{\lambda}), \; \nabla \mathbf{g}(\mathbf{w}), \; \nabla \Phi(\mathbf{w}), \; \mathbf{g}(\mathbf{w})$$

Compute Newton direction

$$\begin{bmatrix} H & \nabla \mathbf{g} \\ \nabla \mathbf{g}^\top & 0 \end{bmatrix} \begin{bmatrix} \Delta \mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = - \begin{bmatrix} \nabla \Phi \\ \mathbf{g} \end{bmatrix}$$

$$\Delta\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ - \boldsymbol{\lambda}$$

Compute Newton step, $t \in \,]0, 1]$

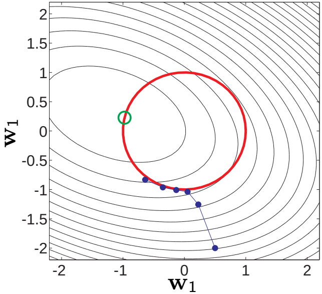
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + t\Delta\boldsymbol{\lambda}$$

return $\mathbf{w}$, $\boldsymbol{\lambda}$

$$\min_{\mathbf{w}} \; \frac{1}{2}\mathbf{w}^\top \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \mathbf{w} + \mathbf{w}^\top \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t.} \;\; g(\mathbf{w}) = \mathbf{w}^\top \mathbf{w} - 1 = 0$$

Guess $\lambda = 0$, step $t = 1$



Newton Iteration for Optimization - Example

Algorithm: Newton method

Input: guess $\mathbf{w}$, $\boldsymbol{\lambda}$

while $\|\nabla\mathcal{L}\|$ *or* $\|\mathbf{g}\| \geq \mathrm{tol}$ **do**

Compute

$$H(\mathbf{w},\boldsymbol{\lambda}),\ \nabla\mathbf{g}(\mathbf{w}),\ \nabla\Phi(\mathbf{w}),\ \mathbf{g}(\mathbf{w})$$

Compute Newton direction

$$\begin{bmatrix} H & \nabla\mathbf{g} \\ \nabla\mathbf{g}^\top & 0 \end{bmatrix} \begin{bmatrix} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = -\begin{bmatrix} \nabla\Phi \\ \mathbf{g} \end{bmatrix}$$

$$\Delta\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ - \boldsymbol{\lambda}$$

Compute Newton step, $t \in]0,1]$

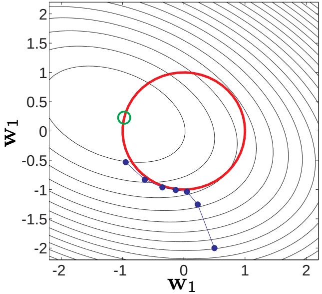
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + t\Delta\boldsymbol{\lambda}$$

return $\mathbf{w}$, $\boldsymbol{\lambda}$

$$\min_{\mathbf{w}} \ \frac{1}{2}\mathbf{w}^\top \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \mathbf{w} + \mathbf{w}^\top \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t.} \ \ g(\mathbf{w}) = \mathbf{w}^\top\mathbf{w} - 1 = 0$$

Guess $\lambda = 0$, step $t = 1$



Newton Iteration for Optimization - Example

Algorithm: Newton method

Input: guess $\mathbf{w}$, $\boldsymbol{\lambda}$

while $\|\nabla\mathcal{L}\|$ *or* $\|\mathbf{g}\| \geq \mathrm{tol}$ **do**

Compute

$$H(\mathbf{w},\boldsymbol{\lambda}),\ \nabla\mathbf{g}(\mathbf{w}),\ \nabla\Phi(\mathbf{w}),\ \mathbf{g}(\mathbf{w})$$

Compute Newton direction

$$\begin{bmatrix} H & \nabla\mathbf{g} \\ \nabla\mathbf{g}^\top & 0 \end{bmatrix}\begin{bmatrix} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = -\begin{bmatrix} \nabla\Phi \\ \mathbf{g} \end{bmatrix}$$

$$\Delta\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ - \boldsymbol{\lambda}$$

Compute Newton step, $t \in]0,1]$

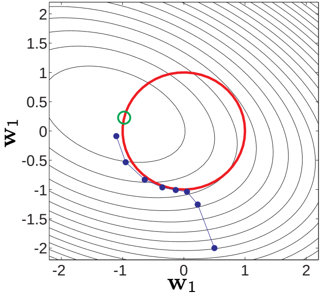
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + t\Delta\boldsymbol{\lambda}$$

return $\mathbf{w}$, $\boldsymbol{\lambda}$

$$\min_{\mathbf{w}}\ \frac{1}{2}\mathbf{w}^\top\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}\mathbf{w} + \mathbf{w}^\top\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t.}\ \ g(\mathbf{w}) = \mathbf{w}^\top\mathbf{w} - 1 = 0$$

Guess $\lambda = 0$, step $t = 1$



Newton Iteration for Optimization - Example

Algorithm: Newton method

Input: guess $\mathbf{w}$, $\boldsymbol{\lambda}$

while $\|\nabla\mathcal{L}\|$ *or* $\|\mathbf{g}\| \geq \mathrm{tol}$ **do**

Compute

$$H(\mathbf{w},\boldsymbol{\lambda}),\ \nabla\mathbf{g}(\mathbf{w}),\ \nabla\Phi(\mathbf{w}),\ \mathbf{g}(\mathbf{w})$$

Compute Newton direction

$$\begin{bmatrix} H & \nabla\mathbf{g} \\ \nabla\mathbf{g}^\top & 0 \end{bmatrix}\begin{bmatrix} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = -\begin{bmatrix} \nabla\Phi \\ \mathbf{g} \end{bmatrix}$$

$$\Delta\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ - \boldsymbol{\lambda}$$

Compute Newton step, $t \in]0,1]$

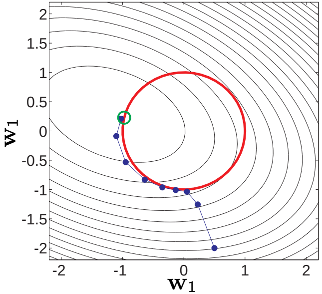
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + t\Delta\boldsymbol{\lambda}$$

return $\mathbf{w}$, $\boldsymbol{\lambda}$

$$\min_{\mathbf{w}} \ \frac{1}{2}\mathbf{w}^\top \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}\mathbf{w} + \mathbf{w}^\top \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t. } g(\mathbf{w}) = \mathbf{w}^\top\mathbf{w} - 1 = 0$$

Guess $\lambda = 0$, step $t = 1$



Newton Iteration for Optimization - Example

Algorithm: Newton method

Input: guess $\mathbf{w}$, $\boldsymbol{\lambda}$

while $\|\nabla\mathcal{L}\|$ *or* $\|\mathbf{g}\| \geq \text{tol}$ **do**

Compute

$$H(\mathbf{w},\boldsymbol{\lambda}),\ \nabla\mathbf{g}(\mathbf{w}),\ \nabla\Phi(\mathbf{w}),\ \mathbf{g}(\mathbf{w})$$

Compute Newton direction

$$\begin{bmatrix} H & \nabla\mathbf{g} \\ \nabla\mathbf{g}^\top & 0 \end{bmatrix}\begin{bmatrix} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = -\begin{bmatrix} \nabla\Phi \\ \mathbf{g} \end{bmatrix}$$

$$\Delta\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ - \boldsymbol{\lambda}$$

Compute Newton step, $t \in \,]0,1]$

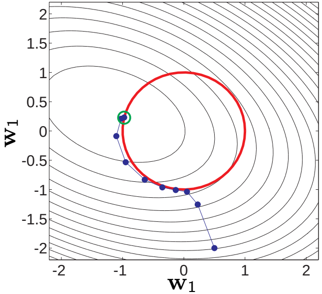
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + t\Delta\boldsymbol{\lambda}$$

return $\mathbf{w}$, $\boldsymbol{\lambda}$

$$\min_{\mathbf{w}} \ \frac{1}{2}\mathbf{w}^\top \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}\mathbf{w} + \mathbf{w}^\top \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t. } g(\mathbf{w}) = \mathbf{w}^\top\mathbf{w} - 1 = 0$$

Guess $\lambda = 0$, step $t = 1$



Newton Iteration for Optimization - Example

Algorithm: Newton method

Input: guess $\mathbf{w}$, $\boldsymbol{\lambda}$

while $\|\nabla\mathcal{L}\|$ *or* $\|\mathbf{g}\| \geq \text{tol}$ **do**

Compute

$$H(\mathbf{w}, \boldsymbol{\lambda}),\ \nabla\mathbf{g}(\mathbf{w}),\ \nabla\Phi(\mathbf{w}),\ \mathbf{g}(\mathbf{w})$$

Compute Newton direction

$$\begin{bmatrix} H & \nabla\mathbf{g} \\ \nabla\mathbf{g}^\top & 0 \end{bmatrix} \begin{bmatrix} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = -\begin{bmatrix} \nabla\Phi \\ \mathbf{g} \end{bmatrix}$$

$$\Delta\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ - \boldsymbol{\lambda}$$

Compute Newton step, $t \in]0, 1]$

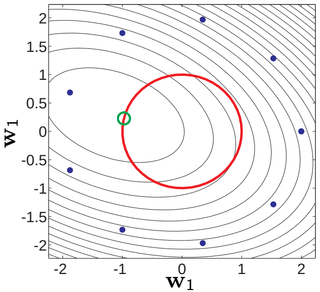
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + t\Delta\boldsymbol{\lambda}$$

return $\mathbf{w}$, $\boldsymbol{\lambda}$

$$\min_{\mathbf{w}} \ \frac{1}{2}\mathbf{w}^\top \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \mathbf{w} + \mathbf{w}^\top \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t. } g(\mathbf{w}) = \mathbf{w}^\top\mathbf{w} - 1 = 0$$

Guess $\lambda = 0$, step $t = 1$



Newton Iteration for Optimization - Example

Algorithm: Newton method

Input: guess $\mathbf{w}$, $\boldsymbol{\lambda}$

while $\|\nabla\mathcal{L}\|$ *or* $\|\mathbf{g}\| \geq \text{tol}$ **do**

Compute

$$H(\mathbf{w}, \boldsymbol{\lambda}),\ \nabla\mathbf{g}(\mathbf{w}),\ \nabla\Phi(\mathbf{w}),\ \mathbf{g}(\mathbf{w})$$

Compute Newton direction

$$\begin{bmatrix} H & \nabla\mathbf{g} \\ \nabla\mathbf{g}^\top & 0 \end{bmatrix} \begin{bmatrix} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = -\begin{bmatrix} \nabla\Phi \\ \mathbf{g} \end{bmatrix}$$

$$\Delta\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ - \boldsymbol{\lambda}$$

Compute Newton step, $t \in]0, 1]$

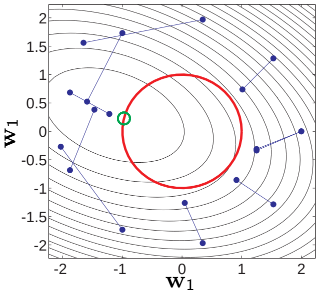
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + t\Delta\boldsymbol{\lambda}$$

return $\mathbf{w}$, $\boldsymbol{\lambda}$

$$\min_{\mathbf{w}} \ \frac{1}{2}\mathbf{w}^\top \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \mathbf{w} + \mathbf{w}^\top \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t.} \ g(\mathbf{w}) = \mathbf{w}^\top\mathbf{w} - 1 = 0$$

Guess $\lambda = 0$, step $t = 1$



Newton Iteration for Optimization - Example

Algorithm: Newton method

Input: guess $\mathbf{w}$, $\boldsymbol{\lambda}$

while $\|\nabla\mathcal{L}\|$ *or* $\|\mathbf{g}\| \geq \mathrm{tol}$ **do**

Compute

$$H\left(\mathbf{w},\boldsymbol{\lambda}\right),\ \nabla\mathbf{g}\left(\mathbf{w}\right),\ \nabla\Phi\left(\mathbf{w}\right),\ \mathbf{g}\left(\mathbf{w}\right)$$

Compute Newton direction

$$\left[\begin{array}{cc} H & \nabla\mathbf{g} \\ \nabla\mathbf{g}^{\top} & 0 \end{array}\right]\left[\begin{array}{c} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^{+} \end{array}\right] = -\left[\begin{array}{c} \nabla\Phi \\ \mathbf{g} \end{array}\right]$$

$$\Delta\boldsymbol{\lambda} = \boldsymbol{\lambda}^{+} - \boldsymbol{\lambda}$$

Compute Newton step, $t \in]0,1]$

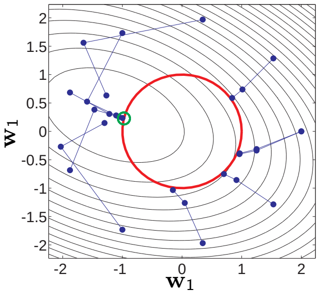
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + t\Delta\boldsymbol{\lambda}$$

return $\mathbf{w}$, $\boldsymbol{\lambda}$

$$\min_{\mathbf{w}}\ \frac{1}{2}\mathbf{w}^{\top}\left[\begin{array}{cc} 2 & 1 \\ 1 & 4 \end{array}\right]\mathbf{w} + \mathbf{w}^{\top}\left[\begin{array}{c} 1 \\ 0 \end{array}\right]$$

$$\text{s.t. } g\left(\mathbf{w}\right) = \mathbf{w}^{\top}\mathbf{w} - 1 = 0$$

Guess $\lambda = 0$, step $t = 1$



Newton Iteration for Optimization - Example

Algorithm: Newton method

Input: guess $\mathbf{w}$, $\boldsymbol{\lambda}$

while $\|\nabla\mathcal{L}\|$ *or* $\|\mathbf{g}\| \geq \text{tol}$ **do**

Compute

$$H(\mathbf{w},\boldsymbol{\lambda}),\ \nabla\mathbf{g}(\mathbf{w}),\ \nabla\Phi(\mathbf{w}),\ \mathbf{g}(\mathbf{w})$$

Compute Newton direction

$$\begin{bmatrix} H & \nabla\mathbf{g} \\ \nabla\mathbf{g}^\top & 0 \end{bmatrix}\begin{bmatrix} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = -\begin{bmatrix} \nabla\Phi \\ \mathbf{g} \end{bmatrix}$$

$$\Delta\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ - \boldsymbol{\lambda}$$

Compute Newton step, $t \in]0,1]$

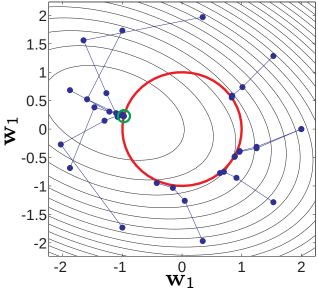
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + t\Delta\boldsymbol{\lambda}$$

return $\mathbf{w}$, $\boldsymbol{\lambda}$

$$\min_{\mathbf{w}}\ \frac{1}{2}\mathbf{w}^\top\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}\mathbf{w} + \mathbf{w}^\top\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t. } g(\mathbf{w}) = \mathbf{w}^\top\mathbf{w} - 1 = 0$$

Guess $\lambda = 0$, step $t = 1$



Newton Iteration for Optimization - Example

Algorithm: Newton method

Input: guess $\mathbf{w}$, $\boldsymbol{\lambda}$

while $\|\nabla\mathcal{L}\|$ *or* $\|\mathbf{g}\| \geq \mathrm{tol}$ **do**

Compute

$$H(\mathbf{w},\boldsymbol{\lambda}),\ \nabla\mathbf{g}(\mathbf{w}),\ \nabla\Phi(\mathbf{w}),\ \mathbf{g}(\mathbf{w})$$

Compute Newton direction

$$\left[\begin{array}{cc} H & \nabla\mathbf{g} \\ \nabla\mathbf{g}^\top & 0 \end{array}\right]\left[\begin{array}{c} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{array}\right] = -\left[\begin{array}{c} \nabla\Phi \\ \mathbf{g} \end{array}\right]$$

$$\Delta\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ - \boldsymbol{\lambda}$$

Compute Newton step, $t \in \,]0,1]$

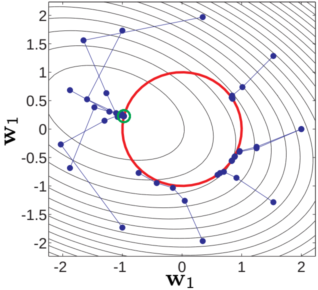
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + t\Delta\boldsymbol{\lambda}$$

return $\mathbf{w}$, $\boldsymbol{\lambda}$

$$\min_{\mathbf{w}} \ \frac{1}{2}\mathbf{w}^\top \left[\begin{array}{cc} 2 & 1 \\ 1 & 4 \end{array}\right]\mathbf{w} + \mathbf{w}^\top\left[\begin{array}{c} 1 \\ 0 \end{array}\right]$$

$$\text{s.t. } g(\mathbf{w}) = \mathbf{w}^\top\mathbf{w} - 1 = 0$$

Guess $\lambda = 0$, step $t = 1$



Newton Iteration for Optimization - Example

Algorithm: Newton method

Input: guess $\mathbf{w}$, $\boldsymbol{\lambda}$

while $\|\nabla\mathcal{L}\|$ *or* $\|\mathbf{g}\| \geq \text{tol}$ **do**

Compute

$$H(\mathbf{w},\boldsymbol{\lambda}),\ \nabla\mathbf{g}(\mathbf{w}),\ \nabla\Phi(\mathbf{w}),\ \mathbf{g}(\mathbf{w})$$

Compute Newton direction

$$\begin{bmatrix} H & \nabla\mathbf{g} \\ \nabla\mathbf{g}^\top & 0 \end{bmatrix}\begin{bmatrix} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = -\begin{bmatrix} \nabla\Phi \\ \mathbf{g} \end{bmatrix}$$

$$\Delta\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ - \boldsymbol{\lambda}$$

Compute Newton step, $t \in]0,1]$

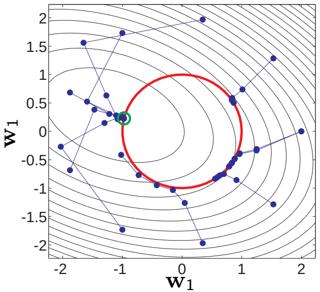
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + t\Delta\boldsymbol{\lambda}$$

return $\mathbf{w}$, $\boldsymbol{\lambda}$

$$\min_{\mathbf{w}} \ \frac{1}{2}\mathbf{w}^\top \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}\mathbf{w} + \mathbf{w}^\top \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t. } g(\mathbf{w}) = \mathbf{w}^\top\mathbf{w} - 1 = 0$$

Guess $\lambda = 0$, step $t = 1$



Newton Iteration for Optimization - Example

Algorithm: Newton method

Input: guess $\mathbf{w}$, $\boldsymbol{\lambda}$

while $\|\nabla\mathcal{L}\|$ *or* $\|\mathbf{g}\| \geq \mathrm{tol}$ **do**

Compute

$$H(\mathbf{w},\boldsymbol{\lambda}),\ \nabla\mathbf{g}(\mathbf{w}),\ \nabla\Phi(\mathbf{w}),\ \mathbf{g}(\mathbf{w})$$

Compute Newton direction

$$\begin{bmatrix} H & \nabla\mathbf{g} \\ \nabla\mathbf{g}^\top & 0 \end{bmatrix}\begin{bmatrix} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = -\begin{bmatrix} \nabla\Phi \\ \mathbf{g} \end{bmatrix}$$

$$\Delta\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ - \boldsymbol{\lambda}$$

Compute Newton step, $t \in]0,1]$

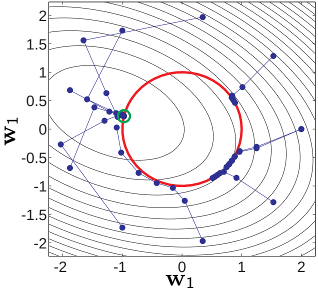
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + t\Delta\boldsymbol{\lambda}$$

return $\mathbf{w}$, $\boldsymbol{\lambda}$

$$\min_{\mathbf{w}}\ \frac{1}{2}\mathbf{w}^\top\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}\mathbf{w} + \mathbf{w}^\top\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t. } g(\mathbf{w}) = \mathbf{w}^\top\mathbf{w} - 1 = 0$$

Guess $\lambda = 0$, step $t = 1$



Newton Iteration for Optimization - Example

Algorithm: Newton method

Input: guess $\mathbf{w}$, $\boldsymbol{\lambda}$

while $\|\nabla\mathcal{L}\|$ *or* $\|\mathbf{g}\| \geq \mathrm{tol}$ **do**

Compute

$$H(\mathbf{w}, \boldsymbol{\lambda}),\ \nabla\mathbf{g}(\mathbf{w}),\ \nabla\Phi(\mathbf{w}),\ \mathbf{g}(\mathbf{w})$$

Compute Newton direction

$$\begin{bmatrix} H & \nabla\mathbf{g} \\ \nabla\mathbf{g}^\top & 0 \end{bmatrix} \begin{bmatrix} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = -\begin{bmatrix} \nabla\Phi \\ \mathbf{g} \end{bmatrix}$$

$$\Delta\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ - \boldsymbol{\lambda}$$

Compute Newton step, $t \in]0, 1]$

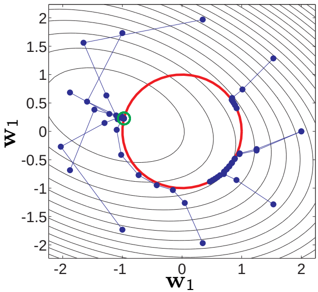
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + t\Delta\boldsymbol{\lambda}$$

return $\mathbf{w}$, $\boldsymbol{\lambda}$

$$\min_{\mathbf{w}} \ \frac{1}{2}\mathbf{w}^\top \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \mathbf{w} + \mathbf{w}^\top \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t.}\ \ g(\mathbf{w}) = \mathbf{w}^\top\mathbf{w} - 1 = 0$$

Guess $\lambda = 0$, step $t = 1$



Newton Iteration for Optimization - Example

Algorithm: Newton method

Input: guess $\mathbf{w}$, $\boldsymbol{\lambda}$

while $\|\nabla\mathcal{L}\|$ *or* $\|\mathbf{g}\| \geq \text{tol}$ **do**

Compute

$$H(\mathbf{w},\boldsymbol{\lambda}),\ \nabla\mathbf{g}(\mathbf{w}),\ \nabla\Phi(\mathbf{w}),\ \mathbf{g}(\mathbf{w})$$

Compute Newton direction

$$\begin{bmatrix} H & \nabla\mathbf{g} \\ \nabla\mathbf{g}^\top & 0 \end{bmatrix}\begin{bmatrix} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = -\begin{bmatrix} \nabla\Phi \\ \mathbf{g} \end{bmatrix}$$

$$\Delta\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ - \boldsymbol{\lambda}$$

Compute Newton step, $t \in \,]0,1]$

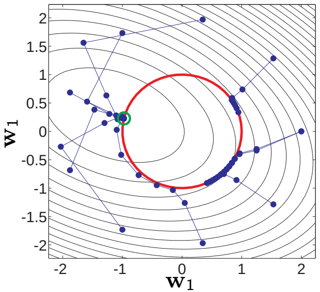
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + t\Delta\boldsymbol{\lambda}$$

return $\mathbf{w}$, $\boldsymbol{\lambda}$

$$\min_{\mathbf{w}} \ \frac{1}{2}\mathbf{w}^\top\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}\mathbf{w} + \mathbf{w}^\top\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t. } g(\mathbf{w}) = \mathbf{w}^\top\mathbf{w} - 1 = 0$$

Guess $\lambda = 0$, step $t = 1$



Newton Iteration for Optimization - Example

Algorithm: Newton method

Input: guess $\mathbf{w}$, $\boldsymbol{\lambda}$

while $\|\nabla\mathcal{L}\|$ *or* $\|\mathbf{g}\| \geq \mathrm{tol}$ **do**

Compute

$$H(\mathbf{w}, \boldsymbol{\lambda}),\ \nabla\mathbf{g}(\mathbf{w}),\ \nabla\Phi(\mathbf{w}),\ \mathbf{g}(\mathbf{w})$$

Compute Newton direction

$$\begin{bmatrix} H & \nabla\mathbf{g} \\ \nabla\mathbf{g}^\top & 0 \end{bmatrix} \begin{bmatrix} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = -\begin{bmatrix} \nabla\Phi \\ \mathbf{g} \end{bmatrix}$$

$$\Delta\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ - \boldsymbol{\lambda}$$

Compute Newton step, $t \in]0, 1]$

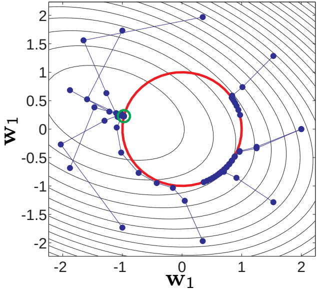
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + t\Delta\boldsymbol{\lambda}$$

return $\mathbf{w}$, $\boldsymbol{\lambda}$

$$\min_{\mathbf{w}} \ \frac{1}{2}\mathbf{w}^\top \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \mathbf{w} + \mathbf{w}^\top \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t. } g(\mathbf{w}) = \mathbf{w}^\top\mathbf{w} - 1 = 0$$

Guess $\lambda = 0$, step $t = 1$



Newton Iteration for Optimization - Example

Algorithm: Newton method

Input: guess $\mathbf{w}$, $\boldsymbol{\lambda}$

while $\|\nabla\mathcal{L}\|$ *or* $\|\mathbf{g}\| \geq \mathrm{tol}$ **do**

Compute

$$H(\mathbf{w},\boldsymbol{\lambda}),\ \nabla\mathbf{g}(\mathbf{w}),\ \nabla\Phi(\mathbf{w}),\ \mathbf{g}(\mathbf{w})$$

Compute Newton direction

$$\begin{bmatrix} H & \nabla\mathbf{g} \\ \nabla\mathbf{g}^\top & 0 \end{bmatrix}\begin{bmatrix} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = -\begin{bmatrix} \nabla\Phi \\ \mathbf{g} \end{bmatrix}$$

$$\Delta\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ - \boldsymbol{\lambda}$$

Compute Newton step, $t \in]0,1]$

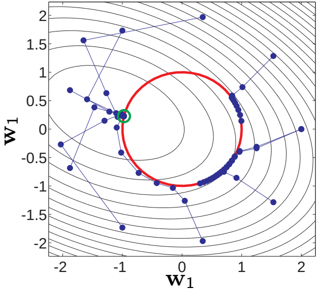
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + t\Delta\boldsymbol{\lambda}$$

return $\mathbf{w}$, $\boldsymbol{\lambda}$

$$\min_{\mathbf{w}} \ \frac{1}{2}\mathbf{w}^\top \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}\mathbf{w} + \mathbf{w}^\top\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t. } g(\mathbf{w}) = \mathbf{w}^\top\mathbf{w} - 1 = 0$$

Guess $\lambda = 0$, step $t = 1$



Newton Iteration for Optimization - Example

Algorithm: Newton method

Input: guess $\mathbf{w}$, $\boldsymbol{\lambda}$

while $\|\nabla\mathcal{L}\|$ *or* $\|\mathbf{g}\| \geq \mathrm{tol}$ **do**

Compute

$$H(\mathbf{w},\boldsymbol{\lambda}),\ \nabla\mathbf{g}(\mathbf{w}),\ \nabla\Phi(\mathbf{w}),\ \mathbf{g}(\mathbf{w})$$

Compute Newton direction

$$\begin{bmatrix} H & \nabla\mathbf{g} \\ \nabla\mathbf{g}^\top & 0 \end{bmatrix}\begin{bmatrix} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = -\begin{bmatrix} \nabla\Phi \\ \mathbf{g} \end{bmatrix}$$

$$\Delta\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ - \boldsymbol{\lambda}$$

Compute Newton step, $t \in]0,1]$

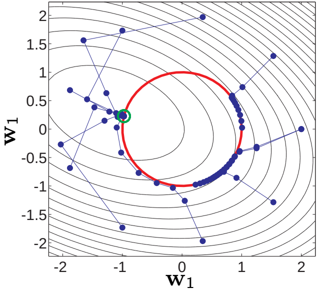
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + t\Delta\boldsymbol{\lambda}$$

return $\mathbf{w}$, $\boldsymbol{\lambda}$

$$\min_{\mathbf{w}}\ \frac{1}{2}\mathbf{w}^\top\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}\mathbf{w} + \mathbf{w}^\top\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t. } g(\mathbf{w}) = \mathbf{w}^\top\mathbf{w} - 1 = 0$$

Guess $\lambda = 0$, step $t = 1$



Newton Iteration for Optimization - Example

Algorithm: Newton method

Input: guess $\mathbf{w}$, $\boldsymbol{\lambda}$

while $\|\nabla\mathcal{L}\|$ *or* $\|\mathbf{g}\| \geq \mathrm{tol}$ **do**

Compute

$$H(\mathbf{w},\boldsymbol{\lambda}),\ \nabla\mathbf{g}(\mathbf{w}),\ \nabla\Phi(\mathbf{w}),\ \mathbf{g}(\mathbf{w})$$

Compute Newton direction

$$\begin{bmatrix} H & \nabla\mathbf{g} \\ \nabla\mathbf{g}^\top & 0 \end{bmatrix}\begin{bmatrix} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = -\begin{bmatrix} \nabla\Phi \\ \mathbf{g} \end{bmatrix}$$

$$\Delta\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ - \boldsymbol{\lambda}$$

Compute Newton step, $t \in]0,1]$

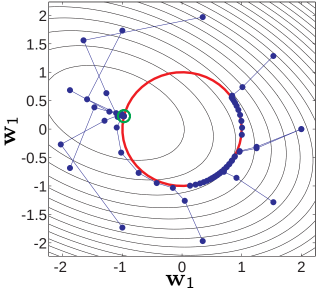
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + t\Delta\boldsymbol{\lambda}$$

return $\mathbf{w}$, $\boldsymbol{\lambda}$

$$\min_{\mathbf{w}}\ \frac{1}{2}\mathbf{w}^\top\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}\mathbf{w} + \mathbf{w}^\top\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t. } g(\mathbf{w}) = \mathbf{w}^\top\mathbf{w} - 1 = 0$$

Guess $\lambda = 0$, step $t = 1$



Newton Iteration for Optimization - Example

Algorithm: Newton method

Input: guess $\mathbf{w}$, $\boldsymbol{\lambda}$

while $\|\nabla\mathcal{L}\|$ *or* $\|\mathbf{g}\| \geq \mathrm{tol}$ **do**

Compute

$$H(\mathbf{w}, \boldsymbol{\lambda}),\ \nabla\mathbf{g}(\mathbf{w}),\ \nabla\Phi(\mathbf{w}),\ \mathbf{g}(\mathbf{w})$$

Compute Newton direction

$$\begin{bmatrix} H & \nabla\mathbf{g} \\ \nabla\mathbf{g}^\top & 0 \end{bmatrix} \begin{bmatrix} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = -\begin{bmatrix} \nabla\Phi \\ \mathbf{g} \end{bmatrix}$$

$$\Delta\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ - \boldsymbol{\lambda}$$

Compute Newton step, $t \in]0, 1]$

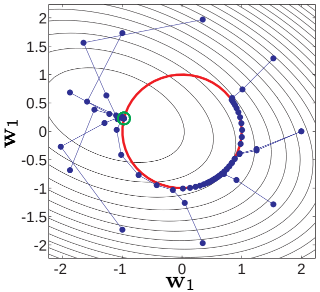
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + t\Delta\boldsymbol{\lambda}$$

return $\mathbf{w}$, $\boldsymbol{\lambda}$

$$\min_{\mathbf{w}} \ \frac{1}{2}\mathbf{w}^\top \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \mathbf{w} + \mathbf{w}^\top \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t. } g(\mathbf{w}) = \mathbf{w}^\top\mathbf{w} - 1 = 0$$

Guess $\lambda = 0$, step $t = 1$



Newton Iteration for Optimization - Example

Algorithm: Newton method

Input: guess $\mathbf{w}$, $\boldsymbol{\lambda}$

while $\|\nabla\mathcal{L}\|$ *or* $\|\mathbf{g}\| \geq \mathrm{tol}$ **do**

Compute

$$H(\mathbf{w}, \boldsymbol{\lambda}),\ \nabla\mathbf{g}(\mathbf{w}),\ \nabla\Phi(\mathbf{w}),\ \mathbf{g}(\mathbf{w})$$

Compute Newton direction

$$\begin{bmatrix} H & \nabla\mathbf{g} \\ \nabla\mathbf{g}^\top & 0 \end{bmatrix} \begin{bmatrix} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = - \begin{bmatrix} \nabla\Phi \\ \mathbf{g} \end{bmatrix}$$

$$\Delta\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ - \boldsymbol{\lambda}$$

Compute Newton step, $t \in \,]0, 1]$

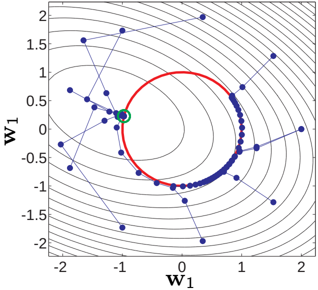
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + t\Delta\boldsymbol{\lambda}$$

return $\mathbf{w}$, $\boldsymbol{\lambda}$

$$\min_{\mathbf{w}} \ \frac{1}{2}\mathbf{w}^\top \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \mathbf{w} + \mathbf{w}^\top \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t. } g(\mathbf{w}) = \mathbf{w}^\top\mathbf{w} - 1 = 0$$

Guess $\lambda = 0$, step $t = 1$



Newton Iteration for Optimization - Example

Algorithm: Newton method

Input: guess $\mathbf{w}$, $\boldsymbol{\lambda}$

while $\|\nabla\mathcal{L}\|$ *or* $\|\mathbf{g}\| \geq \text{tol}$ **do**

Compute

$$H\left(\mathbf{w}, \boldsymbol{\lambda}\right), \nabla\mathbf{g}\left(\mathbf{w}\right), \nabla\Phi\left(\mathbf{w}\right), \mathbf{g}\left(\mathbf{w}\right)$$

Compute Newton direction

$$\left[\begin{array}{cc} H & \nabla\mathbf{g} \\ \nabla\mathbf{g}^\top & 0 \end{array}\right]\left[\begin{array}{c} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{array}\right] = -\left[\begin{array}{c} \nabla\Phi \\ \mathbf{g} \end{array}\right]$$

$$\Delta\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ - \boldsymbol{\lambda}$$

Compute Newton step, $t \in]0, 1]$

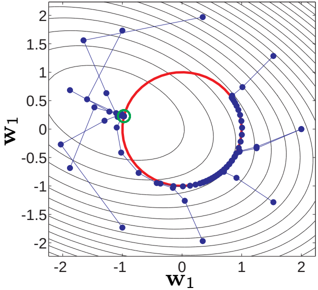
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + t\Delta\boldsymbol{\lambda}$$

return $\mathbf{w}$, $\boldsymbol{\lambda}$

$$\min_{\mathbf{w}} \frac{1}{2}\mathbf{w}^\top \left[\begin{array}{cc} 2 & 1 \\ 1 & 4 \end{array}\right]\mathbf{w} + \mathbf{w}^\top \left[\begin{array}{c} 1 \\ 0 \end{array}\right]$$

$$\text{s.t. } g\left(\mathbf{w}\right) = \mathbf{w}^\top\mathbf{w} - 1 = 0$$

Guess $\lambda = 0$, step $t = 1$



Newton Iteration for Optimization - Example

Algorithm: Newton method

Input: guess $\mathbf{w}$, $\boldsymbol{\lambda}$

while $\|\nabla\mathcal{L}\|$ *or* $\|\mathbf{g}\| \geq \mathrm{tol}$ **do**

Compute

$$H(\mathbf{w},\boldsymbol{\lambda}),\ \nabla\mathbf{g}(\mathbf{w}),\ \nabla\Phi(\mathbf{w}),\ \mathbf{g}(\mathbf{w})$$

Compute Newton direction

$$\begin{bmatrix} H & \nabla\mathbf{g} \\ \nabla\mathbf{g}^\top & 0 \end{bmatrix}\begin{bmatrix} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = -\begin{bmatrix} \nabla\Phi \\ \mathbf{g} \end{bmatrix}$$

$$\Delta\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ - \boldsymbol{\lambda}$$

Compute Newton step, $t \in]0,1]$

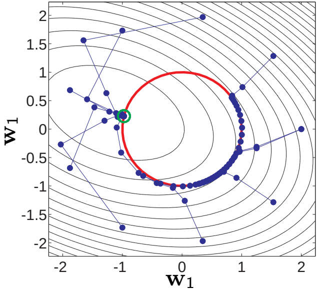
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + t\Delta\boldsymbol{\lambda}$$

return $\mathbf{w}$, $\boldsymbol{\lambda}$

$$\min_{\mathbf{w}} \ \frac{1}{2}\mathbf{w}^\top\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}\mathbf{w} + \mathbf{w}^\top\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t. } g(\mathbf{w}) = \mathbf{w}^\top\mathbf{w} - 1 = 0$$

Guess $\lambda = 0$, step $t = 1$



Newton Iteration for Optimization - Example

Algorithm: Newton method

Input: guess $\mathbf{w}$, $\boldsymbol{\lambda}$

while $\|\nabla\mathcal{L}\|$ *or* $\|\mathbf{g}\| \geq \mathrm{tol}$ **do**

Compute

$$H(\mathbf{w},\boldsymbol{\lambda}),\ \nabla\mathbf{g}(\mathbf{w}),\ \nabla\Phi(\mathbf{w}),\ \mathbf{g}(\mathbf{w})$$

Compute Newton direction

$$\begin{bmatrix} H & \nabla\mathbf{g} \\ \nabla\mathbf{g}^\top & 0 \end{bmatrix}\begin{bmatrix} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = -\begin{bmatrix} \nabla\Phi \\ \mathbf{g} \end{bmatrix}$$

$$\Delta\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ - \boldsymbol{\lambda}$$

Compute Newton step, $t \in]0,1]$

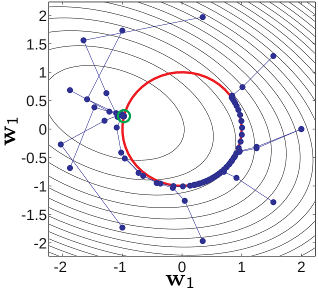
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + t\Delta\boldsymbol{\lambda}$$

return $\mathbf{w}$, $\boldsymbol{\lambda}$

$$\min_{\mathbf{w}} \ \frac{1}{2}\mathbf{w}^\top \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}\mathbf{w} + \mathbf{w}^\top \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t. } g(\mathbf{w}) = \mathbf{w}^\top\mathbf{w} - 1 = 0$$

Guess $\lambda = 0$, step $t = 1$



Newton Iteration for Optimization - Example

Algorithm: Newton method

Input: guess $\mathbf{w}$, $\boldsymbol{\lambda}$

while $\|\nabla\mathcal{L}\|$ *or* $\|\mathbf{g}\| \geq \text{tol}$ **do**

Compute

$$H(\mathbf{w}, \boldsymbol{\lambda}),\ \nabla\mathbf{g}(\mathbf{w}),\ \nabla\Phi(\mathbf{w}),\ \mathbf{g}(\mathbf{w})$$

Compute Newton direction

$$\begin{bmatrix} H & \nabla\mathbf{g} \\ \nabla\mathbf{g}^\top & 0 \end{bmatrix} \begin{bmatrix} \Delta\mathbf{w} \\ \boldsymbol{\lambda}^+ \end{bmatrix} = -\begin{bmatrix} \nabla\Phi \\ \mathbf{g} \end{bmatrix}$$

$$\Delta\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ - \boldsymbol{\lambda}$$

Compute Newton step, $t \in]0, 1]$

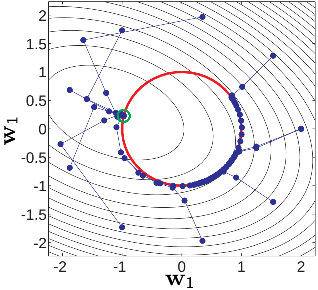
$$\mathbf{w} \leftarrow \mathbf{w} + t\Delta\mathbf{w}, \quad \boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + t\Delta\boldsymbol{\lambda}$$

return $\mathbf{w}$, $\boldsymbol{\lambda}$

$$\min_{\mathbf{w}} \ \frac{1}{2}\mathbf{w}^\top \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \mathbf{w} + \mathbf{w}^\top \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{s.t. } g(\mathbf{w}) = \mathbf{w}^\top\mathbf{w} - 1 = 0$$

Guess $\lambda = 0$, step $t = 1$



Your initial guess matters !!