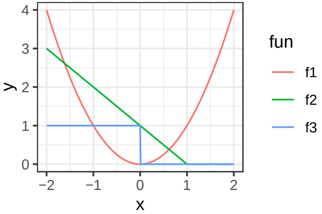
Optimization

Taylor series



Learning goals

- TODO
- TODO

TAYLOR SERIES (MULTIVARIATE)

The Taylor series at point $\tilde{\boldsymbol{x}}$ is

- first order:

$$f(\boldsymbol{x}) = f(\tilde{\boldsymbol{x}}) + \nabla f(\tilde{\boldsymbol{x}})^\top (\mathbf{x} - \tilde{\boldsymbol{x}}) + \mathcal{O}(\|\mathbf{x} - \tilde{\boldsymbol{x}}\|^2)$$

- second order:

$$f(\boldsymbol{x}) = f(\tilde{\boldsymbol{x}}) + \nabla f(\tilde{\boldsymbol{x}})^\top (\mathbf{x} - \tilde{\boldsymbol{x}}) + \frac{1}{2}(\mathbf{x} - \tilde{\boldsymbol{x}})^\top \nabla^2 f(\tilde{\boldsymbol{x}})(\mathbf{x} - \tilde{\boldsymbol{x}}) + \mathcal{O}(\|\mathbf{x} - \tilde{\boldsymbol{x}}\|^3)$$

Note: The order of the error of the taylor approximation is

- smaller at points $\mathbf{x}$ close to $\tilde{\boldsymbol{x}}$
- smaller for the higher the order of the taylor approximation (because higher order approximations give us more flexibility)

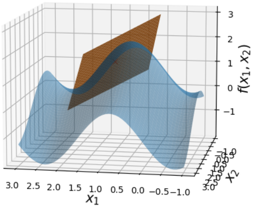
The n^{th} order taylor series is the best n^{th} order approximation to $f(\mathbf{x})$ near $\tilde{\boldsymbol{x}}$.

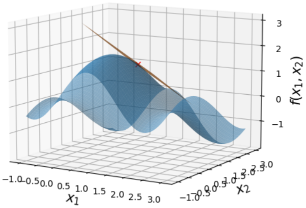
TAYLOR SERIES (MULTIVARIATE)

Example: We consider the function $f(\mathbf{x}) = \sin(2x_1) + \cos(x_2)$.

The gradient is $\nabla f(\mathbf{x}) = (2\cos(2x_1), -\sin(x_2))^\top$. With this, the resulting first order Taylor approximation in $\tilde{\boldsymbol{x}} = (1.0, 1.0)$ is

$$f(\boldsymbol{x}) \quad \approx \quad T_1(\boldsymbol{x}) = f(1.0, 1.0) + (2\cos(2.0), -\sin(1.0))^T \left(\mathbf{x} - \begin{pmatrix} 1.0 \\ 1.0 \end{pmatrix} \right)$$





TAYLOR SERIES (MULTIVARIATE)

To determine the second order Taylor approximation in $\tilde{\boldsymbol{x}} = (1.0, 1.0)$, we need the corresponding Hessian:

$$\nabla^2 f(\mathbf{x}) = \begin{pmatrix} -4\sin(2x_1) & 0 \\ 0 & -\cos(x_2) \end{pmatrix}$$

and get (together with the linear approximation $T_1(\boldsymbol{x})$):

$$f(\boldsymbol{x}) \quad \approx \quad T_1(\boldsymbol{x}) + \frac{1}{2}\left(\mathbf{x} - \begin{pmatrix} 1.0 \\ 1.0 \end{pmatrix}\right)^\top \begin{pmatrix} -4\sin(2.0) & 0 \\ 0 & -\cos(1.0) \end{pmatrix} \left(\mathbf{x} - \begin{pmatrix} 1.0 \\ 1.0 \end{pmatrix}\right)$$