Overcomplete Dictionaries

Ekaterina Lomakina

May 13, 2011

Minimizing the Residual

Atom selection at iteration t:

$$d^*(t) = \arg\max_d \left|\left\langle \mathbf{r}^t, \mathbf{u}_d \right\rangle\right|$$

Proof for first iteration:

- Project $\mathbf{r}^0 = \mathbf{x}$ on atom $\mathbf{u}_d$, to get

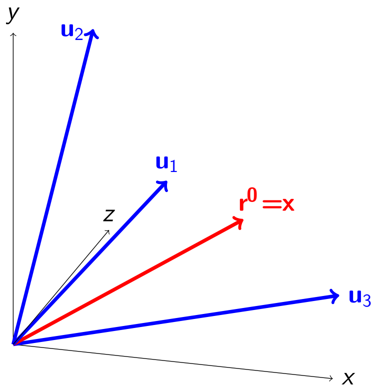
$$\mathbf{x} = \langle \mathbf{x}, \mathbf{u}_d \rangle \mathbf{u}_d + \mathbf{r}^1$$

- Since $\mathbf{r}^1$ is orthogonal to $\mathbf{u}_d$, and $\mathbf{u}_d^\top \mathbf{u}_d = 1$,

$$\|\mathbf{x}\|_2^2 = |\langle \mathbf{x}, \mathbf{u}_d \rangle|^2 + \left\|\mathbf{r}^1\right\|_2^2$$

- Therefore, $\left\|\mathbf{r}^1\right\|_2^2$ is minimized by maximizing $\left|\left\langle \mathbf{r}^0, \mathbf{u}_d \right\rangle\right|^2$.

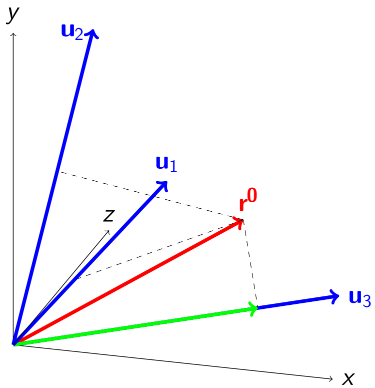
MP Example



Bach et al. (2009)

$$\mathbf{z} = (0, 0, 0)$$

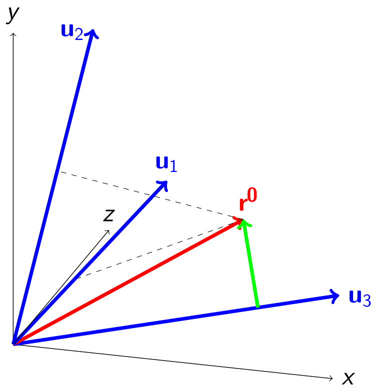
MP Example



Bach et al. (2009)

$$\mathbf{z} = (0, 0, 0)$$

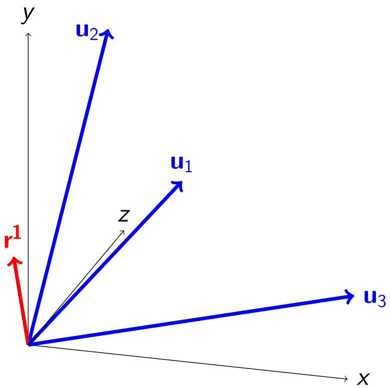
MP Example



Bach et al. (2009)

$$\mathbf{z} = (0, 0, 0.75)$$

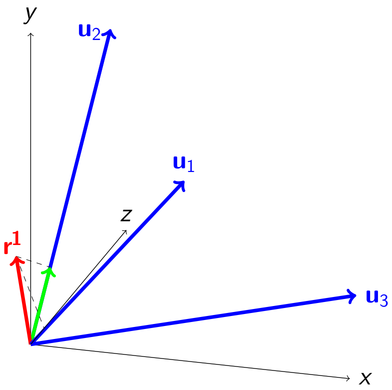
MP Example



Bach et al. (2009)

$$\mathbf{z} = (0, 0, 0.75)$$

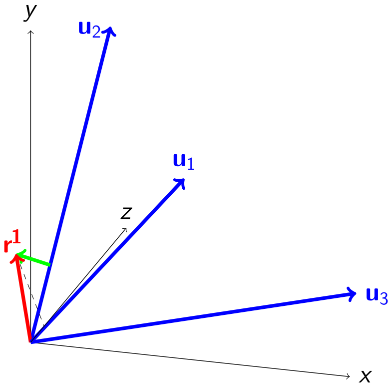
MP Example



Bach et al. (2009)

$$\mathbf{z} = (0, 0, 0.75)$$

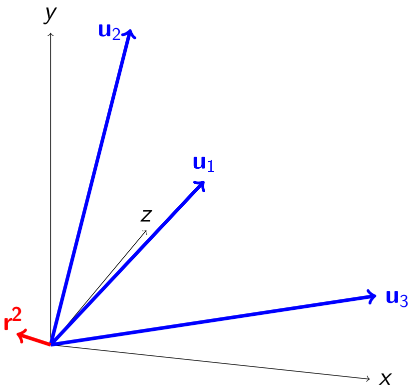
MP Example



Bach et al. (2009)

$$\mathbf{z} = (0, 0.24, 0.75)$$

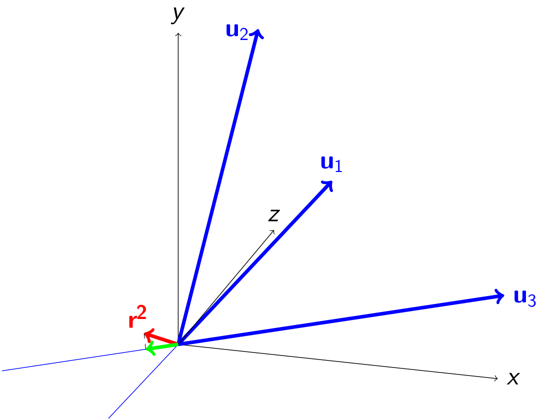
MP Example



Bach et al. (2009)

$$\mathbf{z} = (0, 0.24, 0.75)$$

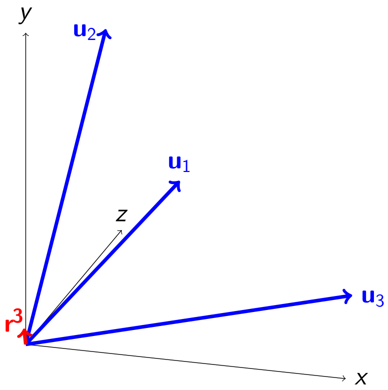
MP Example



Bach et al. (2009)

$$\mathbf{z} = (0, 0.24, 0.65)$$

MP Example



Bach et al. (2009)

$$\mathbf{z} = (0, 0.24, 0.65)$$

Pen&Pencil - I

$$\mathbf{x} = \left\langle \mathbf{x}, \mathbf{u}_{d(1)} \right\rangle \mathbf{u}_{d(1)} + \mathbf{r}^1$$

$\mathbf{r}^1$ is orthogonal to $\mathbf{u}_{d(1)}$

For the next step we have

$$\mathbf{r}^1 = \left\langle \mathbf{r}^1, \mathbf{u}_{d(2)} \right\rangle \mathbf{u}_{d(2)} + \mathbf{r}^2$$

$\mathbf{r}^2$ is orthogonal to $\mathbf{u}_{d(2)}$

Question: Is $\mathbf{r}^2$ orthogonal to $\mathbf{u}_{d(1)}$? When is it true?

Pen&Pencil - II

This is projection given that $\langle \mathbf{u}_d, \mathbf{u}_d \rangle = 1$ for evey d:

$$\mathbf{x} = \left\langle \mathbf{x}, \mathbf{u}_{d(1)} \right\rangle \mathbf{u}_{d(1)} + \mathbf{r}^1$$

$\mathbf{r}^1$ is orthogonal to $\mathbf{u}_{d(1)}$

Question1: What is different when $\langle \mathbf{u}_d, \mathbf{u}_d \rangle \neq 1$?

Question2: Given $\langle \mathbf{u}_d, \mathbf{u}_d \rangle \neq 1$ what should be maxmized to minimize $\left\|\mathbf{r}^1\right\|_2^2$ in $\mathbf{x} = \langle \mathbf{x}, \mathbf{u}_d \rangle \mathbf{u}_d + \mathbf{r}^1$?

MP Algorithm

Objective:

$$\mathbf{z}^* = \arg\min_{\mathbf{z}} \|\mathbf{x} - \mathbf{U}\mathbf{z}\|_2$$
$$\text{s.t.} \quad \|\mathbf{z}\|_0 \leq K$$

Algorithm:

1: $\mathbf{z} \leftarrow \mathbf{0}$, $\mathbf{r} \leftarrow \mathbf{x}$
2: **while** $\|\mathbf{z}\|_0 < K$ **do**
3: Select atom with maximum absolute correlation to residual:

$$d^* \leftarrow \arg\max_d \left|\mathbf{u}_d^\top \mathbf{r}\right|$$

4: Update coefficient vector and residual:

$$z_{d^*} \leftarrow z_{d^*} + \mathbf{u}_{d^*}^\top \mathbf{r}$$

$$\mathbf{r} \leftarrow \mathbf{r} - \left(\mathbf{u}_{d^*}^\top \mathbf{r}\right) \mathbf{u}_{d^*}$$

5: **end while**