1.5 The conjugate gradient method

1.5.1 Motivation and Krylov subspaces

The *conjugate gradient (CG) method* is a very effective algorithm for solving the SPD system $Ax = b$. It is a perfect example of a *Krylov subspace method*, which will be discussed for general systems (not necessarily SPD) in Ch. 1.6. The inverse of the matrix A can be described as a polynomial in A of degree less than n by using the *Cayley-Hamilton theorem*:

Theorem 1.12. *Let $A \in \mathbb{C}^{n\times n}$ be a matrix with characteristic polynomial $\chi(z)$. Then, the matrix A satisifies its own characteristic equation, i.e. $\chi(A) = 0$.* □

In $\mathbb{C}$ we may write χ as

$$\chi(z) = c_n z^n + c_{n-1} z^{n-1} + \ldots + c_1 z + c_0. \tag{47}$$

Theorem 1.12 states that

$$\chi(A) = c_n A^n + c_{n-1} A^{n-1} + \ldots + c_1 A + c_0 I = 0 \tag{48}$$

is satisified as a matrix equation. A direct consequence of (48) for a nonsingular matrix $A \in \mathbb{R}^{n\times n}$ is a representation of the inverse A^{-1} as a matrix polynomial in A of degree less than n, i.e. by multiplying (48) with A^{-1} and dividing by $c_0 = \det(A) \neq 0$ (A is nonsingular) we get

$$A^{-1} = -\frac{c_n}{c_0} A^{n-1} - \frac{c_{n-1}}{c_0} A^{n-2} - \ldots - \frac{c_1}{c_0} I. \tag{49}$$

Now note, that the solution x_* of $Ax = b$ can be written in the form

$$x_* = x_0 + A^{-1} r_0 \tag{50}$$

with the residual $r_0 = b - Ax_0$. The term $A^{-1}r_0$ can be constructed by using the Cayley-Hamilton theorem

$$A^{-1} r_0 = (d_{n-1} A^{n-1} + d_{n-2} A^{n-2} + \ldots + d_1 I) r_0, \tag{51}$$

where the coefficients are determined by the characteristic polynomial. This means that

$$A^{-1} r_0 \in \mathcal{K}_n := \text{span}\{r_0, Ar_0, \ldots, A^{n-1} r_0\}, \tag{52}$$

and hence, the solution x_* satisfies

$$x_* \in x_0 + \mathcal{K}_n. \tag{53}$$

Definition 1.13 (Krylov subspace). *In general, we define for $m \geq 1$ the m-th Krylov space of A w.r.t. the initial residual r_0 as (compare with [1])*

$$\mathcal{K}_m = \mathcal{K}_m(A, r_0) = \mathit{span}\{r_0, Ar_0, A^2 r_0, \ldots, A^{m-1} r_0\}, \tag{54}$$

which has dimension less or equal m.[11]

As will be discussed in Ch. 1.6, Krylov subspace methods try to approximate the solution x_* by

$$x_* \approx x_m \in x_0 + \mathcal{K}_m, \tag{55}$$

with $m \ll n$.

The *conjugate gradient method* is the historically earliest example of such a method, developed for SPD systems by Hestenes, Stiefel and Lanczos in the $50's$. It minimizes the error $\varepsilon_k = x_* - x_k$ in the energy norm while constructing the iterates $x_k \in x_0 + \mathcal{K}_k(A, r_0)$, i.e.

$$x_k = \underset{x \in x_0 + \mathcal{K}_k(A,r_0)}{\arg\min} \|x - x_*\|_A = \underset{x \in x_0 + \mathcal{K}_k(A,r_0)}{\arg\min} \Phi(x), \tag{56}$$

see (33) for the latter equality.

[11] Note, that $x_* = A^{-1} b \in \mathcal{K}_n(A, b)$ is independent of the starting value. Therefore, some authors define the m-th Krylov space as $\mathcal{K}_m(A, b)$.

1.5.2 Conjugate direction methods

The conjugate gradient method is a so-called *conjugate direction method* for minimizing the quadratic functional (31).

Definition 1.14 (Conjugate direction methods). *In a conjugate direction method the new iterates are determined by*

$$x_{k+1} = x_k + \alpha_k p_k, \tag{57}$$

where the set of search directions $\{p_0, p_1, \dots, p_k\}$ are conjugate w.r.t. the SPD matrix A, that is, orthogonal w.r.t. the energy product, i.e., [12]

$$p_i^T A p_j = 0, \quad \text{for all } i \neq j. \tag{58}$$

The step length α_k in (57) *is the minimizer along p_k (line search - compare with* (36)*) given by*

$$\alpha_k = \frac{r_k^T p_k}{p_k^T A p_k}. \tag{59}$$

The property (58) implies linear independence of the directions.
We can motivate the sense in using conjugate directions in the two dimensional case by the following observation:

Remark 1.15 (Motivation for conjugate directions (in 2d)). Consider the convex minimization problem (31) in $\mathbb{R}^2$. Let $x_0 \in \mathbb{R}^2$ be any initial guess. We choose any descent direction $p_0 \in \mathbb{R}^2$, that is a direction p_0 with $p_0^T \nabla\Phi(x_0) < 0$ [13] Let us define the new approximation as $x_1 = x_0 + \alpha_0 p_0$ with the optimal α_0 as in (59). Then the directions $x_* - x_1$ and p_0 are orthogonal w.r.t. the A-inner product:

$$\langle p_0, x_* - x_1 \rangle_A = p_0^T A(x_* - x_0 - \alpha_0 p_0) = p_0^T r_0 - \alpha_0 \, p_0^T A p_0 = p_0^T r_0 - r_0^T p_0 = 0. \tag{60}$$

This means, convergence in two steps (in the two dimensional case) requires the directions p_0 and p_1 to be conjugate. □

In fact, there holds the following remarkable theorem:

Theorem 1.16. *For any starting vector x_0 the sequence generated by a conjugate direction method* (57) *converges to the solution x_* of the system $Ax = b$, where A is $n \times n$ and SPD, in at most n steps.*

Proof. From the conjugacy of a set $\{p_i\}_{i=0,\dots,n-1}$ one gets linear independence of this set. In particular, there holds $\mathbb{R}^n = \operatorname{span}\{p_0, p_1, \dots, p_{n-1}\}$. This means one can represent the initial error as

$$x_* - x_0 = \sigma_0 p_0 + \sigma_1 p_1 + \cdots + \sigma_{n-1} p_{n-1} \quad \text{for } \sigma_j \in \mathbb{R},\ j = 0, \dots, n-1. \tag{61}$$

Multiplying (61) with $p_k^T A$ for $k \in \{0, \dots, n-1\}$ yields, owing to conjugacy,

$$\sigma_k = \frac{p_k^T A(x_* - x_0)}{p_k^T A p_k}. \tag{62}$$

The proof is done when we show that $\sigma_k = \alpha_k$, the optimally chosen step length from line search (59): A conjugate direction algorithm generates the iterate

$$x_k = x_0 + \alpha_0 p_0 + \alpha_1 p_1 + \cdots + \alpha_{k-1} p_{k-1}.$$

Multiplication with $p_k^T A$ yields (analogue to above) $p_k^T A(x_k - x_0) = 0$ and hence

$$p_k^T A(x_* - x_0) = p_k^T A\big((x_* - x_k) + (x_k - x_0)\big) = p_k^T A(x_* - x_k) = p_k^T (b - A x_k) = p_k^T r_k. \tag{63}$$

[12] The eigenvectors of A symmetric are an example for conjugate directions.
[13] The angle between the steepest descent direction $-\nabla\Phi(x_0)$ at x_0 and the descent direction p_0 is between $-90°$ and $90°$.

Now compare (62) with (59) to see $\sigma_k = \alpha_k$. □

In two dimensions one can also illustrate the special role of conjugate directions by the following [11] interpretation. Choosing conjugate directions corresponds to a coordinate transform where the contour lines of the quadratic function (31) are ellipses which are aligned along the coordinate directions in the transformed system. A solution can therefore be found by successive minimization along the coordinate directions (these are the conjugate directions) in at most two steps, see Fig. 7. In contrast, successive minimization along the coordinate directions in the original system does not give the exact solution in two steps because in general the contour ellipses are no longer aligned along the coordinate directions.

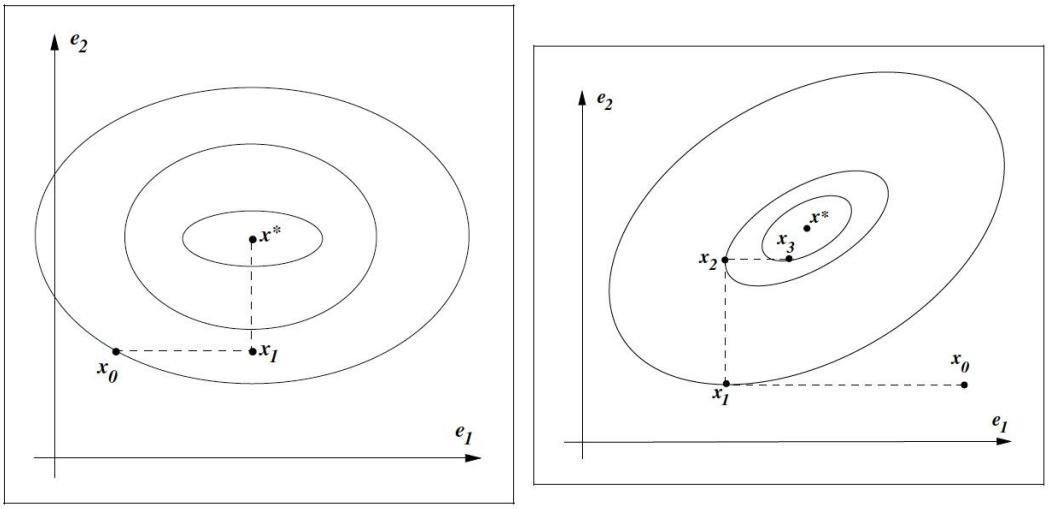

Figure 7: Two dimensional illustration of the convergence process for conjugate direction method (left) and simple successive coordinate direction minimization (right). Taken from [11].

There could be many choices for the set of conjugate directions, e.g. the eigenvectors of a symmetric matrix are pairwise orthogonal and conjugate. For large-scale problems the computation of (all) eigenvectors is too expensive. An alternative would be to modify the Gram-Schmidt orthogonalization process to orthogonality w.r.t. the energy product. But this is also very expensive since it would require to store all previous directions. In contrast, the conjugate gradient method is special and ultimately efficient because it generates the conjugate direction set by only using the previous direction. More precisely, the next direction is chosen as a linear combination of the steepest descent direction and the previous search direction.

1.5.3 The CG method

Definition 1.17 (Conjugate direction and step length in conjugate gradient (CG) method). *The CG method uses the linear combination*

$$p_k = r_k + \beta_k p_{k-1}, \tag{64}$$

to enforce conjugacy of p_k. Hence, the scalar β_k is determined by the requirement that p_k and p_{k-1} are conjugate, which yields

$$\beta_k = -\frac{p_{k-1}^T A r_k}{p_{k-1}^T A p_{k-1}} = -\frac{r_k^T A p_{k-1}}{p_{k-1}^T A p_{k-1}}. \tag{65}$$

The step length α_k is determined by one dimensional minimization along the new direction, see (59)*. The first direction in the CG method is the steepest descent direction.*

There holds the following theorem which shows that the search directions are indeed mutually conjugate and that CG is a Krylov subspace method.

Theorem 1.18. *Let A be SPD. Suppose the k-th iterate of the CG method is not the solution x_* of the SPD system $Ax = b$. Then, there holds*

$$p_k^T A p_i = 0 \quad \textit{for } i = 0, 1, \ldots, k-1, \tag{66}$$

and

$$\textit{span}\{p_0, p_1, \ldots, p_{k-1}\} = \textit{span}\{r_0, Ar_0, \ldots, A^{k-1}r_0\} = \mathcal{K}_k(A, r_0). \tag{67}$$

The CG method converges to x_* in at most n steps.

Ad proof. The proof is by induction with no special technical difficulties. It can be found e.g. in [11]. □

1.5.4 The practical algorithm

There are two more orthogonality relations involved here: It can be shown [11] that for conjugate direction methods there holds

$$r_k^T p_i = 0 \quad \text{for } i = 0, 1, \ldots, k-1 \tag{68}$$

and for the residuals in the CG method [14]

$$r_k^T r_i = 0 \quad \text{for } i = 0, 1, \ldots, k-1. \tag{69}$$

This leads to simplifications for the computation of the coefficients α_k and β_k:

Lemma 1.19. *The coefficients involved in the CG method as in Def. 1.17 can be reformulated in a practically more convenient way as*

$$\beta_{k+1} = -\frac{r_{k+1}^T A p_k}{p_k^T A p_k} = \frac{\|r_{k+1}\|_2^2}{\|r_k\|_2^2} \tag{70}$$

and

$$\alpha_k = \frac{r_k^T p_k}{p_k^T A p_k} = \frac{\|r_k\|_2^2}{\|p_k\|_A^2}. \tag{71}$$

Proof. From the construction rule $p_k = r_k + \beta_k p_{k-1}$ (64) of the conjugate direction we get by left-multiplying with r_k^T and using Eqn. (68) that

$$r_k^T p_k = r_k^T r_k = \|r_k\|_2^2, \tag{72}$$

and hence

$$\alpha_k = \frac{r_k^T p_k}{p_k^T A p_k} = \frac{\|r_k\|_2^2}{\|p_k\|_A^2}. \tag{73}$$

To simplify the expression (65) for β_k, note that from the update rule (57) one gets

$$r_{k+1} - r_k = -\alpha_k A p_k. \tag{74}$$

By left-multiplying (74) with p_k^T and using (68) and relation (72) we get

$$p_k^T A p_k = \frac{1}{\alpha_k} p_k^T r_k = \frac{1}{\alpha_k} \|r_k\|_2^2. \tag{75}$$

Now, left-multiplication of (74) with r_{k+1}^T and using (69) yields

$$r_{k+1}^T A p_k = -\frac{1}{\alpha_k} \|r_{k+1}\|_2^2. \tag{76}$$

[14] Use (68) and $p_i = r_i + \beta_i p_{i-1}$, $i > 1$ and $p_0 = r_0$.

Finally, we get from (75) and (76)

$$\beta_{k+1} = -\frac{r_{k+1}^T A p_k}{p_k^T A p_k} = \frac{\|r_{k+1}\|_2^2}{\|r_k\|_2^2}. \tag{77}$$

□

These simplifications allow to reduce the necessary matrix vector multiplications to only the product Ap_k in each iteration in the CG method. The previous residual is stored and the update formula (74) is used for the new residual. The CG method is summarized in Alg. 4.

Algorithm 4: CG method for solving $Ax = b$, A SPD.

Data: Matrix $A \in \mathbb{R}^{n\times n}$ SPD, r.h.s. $b \in \mathbb{R}^n$, initial vector $x_0 \in \mathbb{R}^n$
Result: $x \in \mathbb{R}^n$ approximate solution
Initialization: Set $r_0 \leftarrow b - Ax_0$, $p_0 \leftarrow r_0$, $k \leftarrow 0$.
while *Convergence criterion not satisfied* **do**

$$\alpha_k \leftarrow \frac{r_k^T r_k}{p_k^T A p_k} \quad \text{see (71)}$$
$$x_{k+1} \leftarrow x_k + \alpha_k p_k \quad \text{see (57)}$$
$$r_{k+1} \leftarrow r_k - \alpha_k A p_k \quad \text{see (74)}$$
$$\beta_{k+1} \leftarrow \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k} \quad \text{see (70)}$$
$$p_{k+1} \leftarrow r_{k+1} + \beta_{k+1} p_k \quad \text{see (64)}$$
$$k \leftarrow k + 1$$

end
$x \leftarrow x_k$

1.5.5 Convergence of the CG method

For conjugate direction methods there holds the following property.

Theorem 1.20 (Expanding subspace minimization; Proof [11]). *Let $x_0 \in \mathbb{R}^n$ and x_k generated by a conjugate direction method (Def 1.14) with conjugate directions p_i, $i = 0, \ldots, k-1$. Then there holds the expanding subspace minimization property, that is*

$$x_k = \underset{\left\{x | x \in x_0 + span\{p_0, \ldots, p_{k-1}\}\right\}}{\arg\min} \tfrac{1}{2} x^T A x - b^T x. \tag{78}$$

□

In addition we know from (67) that for the CG method the expanding subspaces coincide with the Krylov subspaces, i.e.,

$$\text{span}\{p_0, p_1, \ldots, p_{k-1}\} = \text{span}\{r_0, Ar_0, \ldots, A^{k-1} r_0\} = \mathcal{K}_k(A, r_0). \tag{79}$$

The rate of convergence of the CG method can be estimated with the following lemma, which uses the expanding subspace minimization property.

Lemma 1.21. *For the error of the iterates x_k generated by the CG method holds*

$$\|x_* - x_{k+1}\|_A^2 \leq \left(\min_{P_k \in \mathcal{P}_k} \max_{1 \leq i \leq n} \left[1 - \lambda_i P_k(\lambda_i)\right]^2 \right) \|x_* - x_0\|_A^2, \tag{80}$$

where $\mathcal{P}_k$ denotes polynomials of degree smaller or equal k and $\lambda_i > 0$, $i = 1, \ldots, n$ the eigenvalues of the SPD matrix A.

Proof. From (79) we have (with some constants γ_i, $i = 0, \dots, k$)

$$x_{k+1} = x_0 + \gamma_0 r_0 + \gamma_1 A r_0 + \cdots + \gamma_k A^k r_0 = x_0 + P_k^*(A) r_0, \tag{81}$$

for a polynomial $P_k^* \in \mathcal{P}_k$. Now recall from (33) the property

$$\frac{1}{2}\|x_* - x\|_A^2 = \Phi(x) - \Phi(x_*). \tag{82}$$

This, together with the expanding subspace minimization property, shows that the polynomial P_k^* solves the problem

$$\min_{P_k \in \mathcal{P}_k} \|x_* - (x_0 + P_k(A) r_0)\|_A, \tag{83}$$

where, due to $r_0 = A(x_* - x_0)$, further holds

$$\|x_* - x_{k+1}\|_A = \|x_* - (x_0 + P_k(A) r_0)\|_A = \| \left[I - P_k(A) A\right] (x_* - x_0)\|_A. \tag{84}$$

Since A is SPD, we have a orthonormal spectral decomposition $A = V \Lambda V^T$, where we denote the eigenvalues with $0 < \lambda_1 \leq \cdots \leq \lambda_n$ and associated orthonormal eigenvectors with v_i, $i = 1, \dots, n$, which form a basis of $\mathbb{R}^n$. Specifically, we can write the initial error as a linear combination of the eigenbasis, that is $x_* - x_0 = \sum_{i=1}^n \xi_i v_i$ with some coefficients ξ_i, $i = 1, \dots, n$. Eigenvectors of A are also eigenvectors of $P_k(A)$ and we have $P_k(A) v_i = P_k(\lambda_i) v_i$, $i = 1, \dots, n$. We therefore have

$$x_* - x_{k+1} = \sum_{i=1}^n \left[1 - \lambda_i P_k^*(\lambda_i)\right] \xi_i v_i, \tag{85}$$

and due to the orthonormality of the v_i, i.e. $v_i^T v_j = \delta_{ij}$, we get for the respective energy norm of the error

$$\|x_* - x_{k+1}\|_A^2 = \sum_{i=1}^n \lambda_i \left[1 - \lambda_i P_k^*(\lambda_i)\right]^2 \xi_i^2 = \min_{P_k \in \mathcal{P}_k} \sum_{i=1}^n \lambda_i \left[1 - \lambda_i P_k(\lambda_i)\right]^2 \xi_i^2. \tag{86}$$

Extracting the polynomial expression finally leads to

$$\|x_* - x_{k+1}\|_A^2 \leq \min_{P_k \in \mathcal{P}_k} \max_{1 \leq i \leq n} \left[1 - \lambda_i P_k(\lambda_i)\right]^2 \left(\sum_{j=1}^n \lambda_j \xi_j^2 \right) = \min_{P_k \in \mathcal{P}_k} \max_{1 \leq i \leq n} \left[1 - \lambda_i P_k(\lambda_i)\right]^2 \|x_* - x_0\|_A^2, \tag{87}$$

which concludes the proof. □

In some cases it is possible to estimate the nonnegative factor $\min_{P_k \in \mathcal{P}_k} \max_{1 \leq i \leq n} \left[1 - \lambda_i P_k(\lambda_i)\right]^2$ to quantify the rate of convergence. An interesting case occurs if the spectrum of A consists of less than n distinct eigenvalues.

Theorem 1.22. *Let A be SPD with $r \leq n$ distinct eigenvalues. Then the CG method produces the exact solution in at most r iterations.*

Proof. Let us denote the eigenvalues of A with $\lambda_1, \dots, \lambda_n$ taking on the r distinct values $0 < \tau_1 < \cdots < \tau_r$. Consider the polynomial $Q_r(\lambda) \in \mathcal{P}_r$ defined as

$$Q_r(\lambda) = \frac{(-1)^r}{\tau_1 \tau_2 \cdots \tau_r} (\lambda - \tau_1)(\lambda - \tau_2) \cdots (\lambda - \tau_r). \tag{88}$$

We have $Q_r(\lambda_i) = 0$, $i = 1, \dots, n$ and $Q_r(0) = 1$, and hence $1 - Q_r(\lambda) \in \mathcal{P}_r$ with root at $\lambda = 0$. By polynomial division we can define the polynomial with degree $r - 1$

$$\widetilde{P}_{r-1} = \frac{1 - Q_r(\lambda)}{\lambda} \in \mathcal{P}_{r-1}. \tag{89}$$

For the convergence factor in Lemma 1.21 we get with $k = r - 1$

$$0 \leq \min_{P_{r-1} \in \mathcal{P}_{r-1}} \max_{1 \leq i \leq n} \left[1 - \lambda_i P_{r-1}(\lambda_i)\right]^2 \leq \max_{1 \leq i \leq n} \left[1 - \lambda_i \widetilde{P}_{r-1}(\lambda_i)\right]^2 = \max_{1 \leq i \leq n} Q_r(\lambda_i)^2 = 0. \tag{90}$$

Thus, from Lemma 1.21 we get $\|x_* - x_r\|_A^2 = 0$. □

A similar behavior can be expected if A has clustered eigenvalues, see Fig. 8.

There also holds a similar theorem as Th. 1.11 for the SD method, but only the square root of the condition number occurs in the upper bound:

Theorem 1.23 (Proof [1]). *The CG method applied to the SPD system $Ax = b$ yields the following estimate for the error $\varepsilon_k := x_* - x_k$ in the energy norm*

$$\|\varepsilon_k\|_A \leq 2\Big(\frac{\sqrt{\kappa_2(A)} - 1}{\sqrt{\kappa_2(A)} + 1}\Big)^k \|\varepsilon_0\|_A \,. \tag{91}$$

□

Although this is an improvement over Th. 1.11 for the steepest descent method, Th. 1.23 might be still too pessimistic. A major difference over SD is the fact that convergence of CG depends on the whole spectrum of A (not only the extreme eigenvalues). In fact, as Th. 1.22 shows, the CG method converges in at most r steps if A has $r < n$ distinct eigenvalues. Similar behavior occurs if there are clusters of eigenvalues, see Fig. 8.

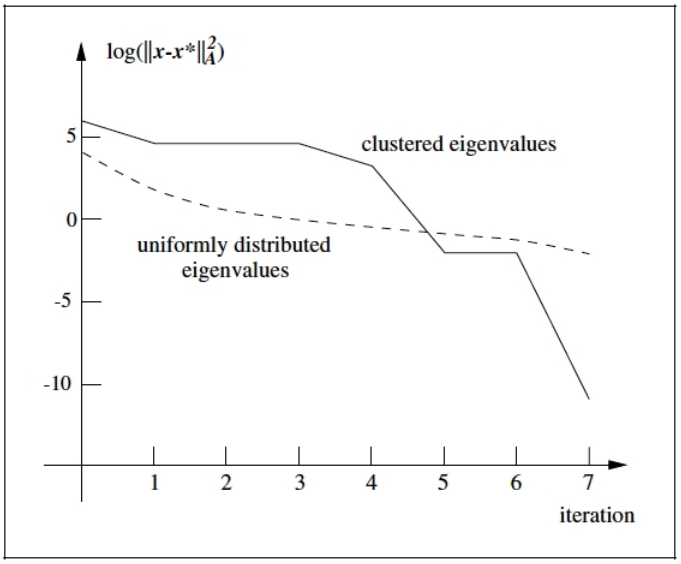

Figure 8: Convergence example of CG for uniformly distributed and clustered spectrum, taken from [11].

1.6 Krylov subspace methods

1.6.1 Krylov subspace projection

We now turn to the general case where $A \in \mathbb{R}^{n \times n}$ is an arbitrary nonsingular matrix.

The basic concept that we will use is a projection property called *Petrov-Galerkin orthogonality*:
Assume we have two subspaces of $\mathbb{R}^n$ that we denote with $\mathcal{K}$ and $\mathcal{L}$. These linear subspaces might be of reduced dimensionality, e.g. $m < n$. Starting from an initial vector $x_0 \in \mathbb{R}^n$ we want to define an